

DISSERTATION

PROTEIN BASED TECHNOLOGIES TO IDENTIFY, STUDY, AND CONTROL
INTRACELLULAR PROCESSES

Submitted by

Virginia Jane Bruce

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

Advisor: Brian McNaughton

Eric Ross
Santiago Di Pietro
Nick Fisk

Copyright by Virginia Bruce 2017

All Rights Reserved

ABSTRACT

PROTEIN BASED TECHNOLOGIES TO IDENTIFY, STUDY, AND CONTROL INTRACELLULAR PROCESSES

Proteins are increasingly used as basic research tools and therapeutics. Their large size, complex structure, and functional group diversity, by virtue of amino acids, often permit recognition of surfaces that challenge small-molecules. Fundamentally, this thesis describes protein based solutions to identifying macromolecules that function inside mammalian cells, enabling visualization and the study of complex biological processes in cellular environments. It also describes the development of engineered polycationic cell-penetrating nanobodies that access the cytosol and thus possibly represent a general solution to intracellularly targeted biologics drug discovery. Collectively, the work described in this thesis reports on: (1) a prostate cancer cell-selective cell penetrating peptide, and its optimization; (2) cell-penetrating nanobodies that access the cytosol of mammalian cells; (3) the use of engineered protein assemblies for bioanalytical reagents and to visualize transcription and translation in mammalian cells, and; (4) a new protein reassembly-based technology to measure cytosolic residence of intracellularly delivered proteins. Technologies and methods described in this work advance the use of proteins in basic science and therapeutically-relevant environments.

ACKNOWLEDGEMENTS

I would like to express sincere gratitude to my research advisor, Prof. Brian McNaughton. More than a mentor, he has been a sponsor of my education and development as a person. I am grateful for the continuous support of my PhD work, for his patience and motivation.

A very special gratitude also goes out to my committee members and CSU Ventures for their support throughout my graduate school experience and providing desired comments and perspectives.

I am indebted to my fellow peers and lab mates. I would like to thank them for their feedback, cooperation and of course friendship. You all have been imperative to helping me grow as a scientist and colleague.

I would like to thank my friends for accepting nothing less than my best. Your outside help and perspective is appreciated more than you can know. Dedication and hard work was learned from my parents and grandparents; I would not be the person I am without their lives as examples. I would like to my thank my little sister, Garnett, for constantly reminding me about effort and grace in every situation and the respect needed for both. Finally, I am grateful to my boyfriend Patrick for always being exactly what I need. Without his support, encouragement and many thoughtful conversations I would not have accomplished as much.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES.....	viii
LIST OF FIGURES	ix
CHAPTER ONE	1
Inside Job: Methods for Delivering Exogenous Proteins to Mammalian Cells.....	1
1.1 Introduction.....	1
1.2 Taking a Step Back: Delivering Nucleic Acids to Mammalian Cells	3
1.2.1 Cationic Lipids and Polymers.....	3
1.2.2 Viral Carriers for the Intracellular Delivery of Protein-Encoding Nucleic Acids	7
1.2.3 Lipid Bilayer Deformation / Physical Insertion.....	8
1.3 Methods for Delivering Exogenous Proteins to Mammalian Cells	10
1.3.1 Cell-Penetrating Peptides.....	11
1.3.2 Cell-Selective Cell-Penetrating Peptides	11
1.3.3 Protein Resurfacing.....	13
1.3.4 The Great Escape: Endosomolytic Peptides	17
1.3.5 Employing Toxins to Deliver Exogenous Proteins.....	19
1.3.6 Membrane – Disruption Based Techniques	22
1.4 Alternative Carriers for Protein Delivery	24
1.5 Conclusion	27
REFERENCES	29

CHAPTER TWO	41
Mutagenesis Modulates the Uptake Efficiency, Cell-Selectivity, and Functional Enzyme	
Delivery of a Protein Transduction Domain.....	41
2.1 Introduction.....	41
2.2 Alanine Scanning of Ypep Illuminates the Importance of Each Residue for PC-3 Cell-Specific Uptake.....	44
2.3 Optimization of Ypep Mutants	45
2.4 Cytotoxicity and Mechanism of Uptake for Ypep Mutants.....	47
2.5 Gly4Asn Ypep Outperforms Tat and Penetration Protein Transduction.....	49
2.6 Mutations Beneficial to Uptake Efficiency Also Increase the Cell-Selectivity of Protein Delivery	50
2.7 Gly4Asn Ypep Mutant Delivers Functional Enzyme to PC-3 Cells	51
2.8 Conclusion	52
REFERENCES	54
CHAPTER THREE	57
Resurfaced Cell-Penetrating Nanobodies: A Potentially General Scaffold for Intracellularly Targeted Protein Discovery	57
3.1 Introduction.....	57
3.2 Polycationic Resurfacing of Three Previously Reported Nanobody Frameworks	60
3.3 Polycationic Resurfacing Does Not Alter Structure, but does Endow Internalization of Mammalian Cells.....	62
3.4 Polycationic Resurfaced Nanobodies Access the Cytosol of Mammalian Cells.....	64
3.5 Polycationic Resurfacing Does Not Alter Nanobody Function and Stability.....	66

3.6 Conclusion	68
REFERENCES	70
CHAPTER FOUR.....	74
Minimalist Antibodies and Mimetics: An Update and Recent Applications.....	74
4.1 Introduction.....	74
4.1.1 Antibodies: Structure, Function, Virtues, and Challenges.....	75
4.2 Antibody Fragments	78
4.2.1 Fab Fragments.....	78
4.2.2 Single-chain Fragment Variables (scFvs).....	80
4.2.3 Minibodies	82
4.2.4 Diabodies	83
4.3 Nanobodies – A Camelid-Derived Scaffold	84
4.3.1 Recent applications of nanobodies.....	86
4.4 Monobodies – A Fibronectin-Derived Scaffold	89
4.4.1 Recent applications of monobodies	90
4.5 Conclusion	92
REFERENCES	94
CHAPTER FIVE	105
Evaluation of Nanobody Conjugates and Protein Fusions as Bioanalytical Reagents	105
5.1 Introduction.....	105
5.2 Enzyme-Linked Immunosorbent Assay (ELISA).....	109
5.3 Flow Cytometry Applications.....	111
5.4 Western Blot Applications.....	114

5.5 Conclusion	116
REFERENCES	117
CHAPTER SIX.....	120
Part I: Intracellular NanoBit Assembly: A Facile Method to Measure Cytosolic Residence of Delivered Proteins and Part II: Innovative Platforms for <i>In Vivo</i> Molecular Tagging and Imaging	120
6.1 Introduction.....	120
6.2 Part I: Intracellular NanoBit Assembly: A Facile Method to Measure Cytosolic Residence of Delivered Proteins	121
6.3 Development of Intracellular NanoBit Assembly	125
6.4 Design of Live-Cell Quantitative NanoBit Luminescence Transduction Assay	126
6.5 Continuing Work – NanoBit.....	131
6.6 Conclusion – NanoBit.....	131
6.7 Part II: Innovate Platforms for <i>In Vivo</i> Molecular Tagging and Imaging	132
6.8 Design of RNA Binders for Imaging Dynamics in Living Cells.....	136
6.9 Design of Protein Binders for Imaging Dynamics in Living Cells.....	142
6.10 Conclusion – Imaging Constructs.....	145
REFERENCES	147
APPENDICES	154
Appendix A: Supplemental Information of Chapter Two	154
Appendix A.1: Chapter Two – Methods.....	154
Appendix A.2: Chapter Two – Proteins Used in this Work	158
Appendix A.3: Chapter Two – Supplemental Information.....	160

Appendix B: Supplemental Information of Chapter Three	167
Appendix B.1: Chapter Two – Methods	167
Appendix B.2: Chapter Three – Proteins Used in this Work.....	172
Appendix B.3: Chapter Three – Supplemental Information.....	174
Appendix C: Supplemental Information of Chapter Five.....	178
Appendix C.1: Chapter Five – Methods	178
Appendix C.2: Chapter Five – Proteins Used in this Work.....	185
Appendix C.3: Chapter Five – Supplemental Information	188
Appendix D: Supplemental Information of Chapter Six	195
Appendix D.1: Chapter Six – Part I - Methods.....	195
Appendix D.2: Chapter Six – Part I: Proteins Used in this Work	199
Appendix D.3: Chapter Six – Part I: Supplemental Information.....	203
Appendix D.4: Chapter Six – Part II - Methods	205
Appendix D.5: Chapter Six – Part II: Proteins Used in this Work	207
Appendix D.6: Chapter Six – Part II: Supplemental Information	209
ABBREVIATIONS	210

LIST OF TABLES

Table 6. 1 Proteins to be assayed for delivery and cytosolic residence.....	128
Table 6. 2 Comparison of MS2 RNA binding protein with U1A and TBP6.7.....	138
Table S5. 1 Flow Cytometry Data – Bacteria.....	190
Table S5. 2 Flow Cytometry Data- Yeast.....	191

LIST OF FIGURES

Figure 1. 1 Common Methods for Achieving Intracellular Delivery of Protein-Encoding Nucleic Acids	4
Figure 1. 2 Uptake Mechanism for Nucleic Acid Delivery with Cationic Lipids.	5
Figure 1. 3 Ionizable Lipid Nanocarriers.....	6
Figure 1. 4 Cancer Cell Activated CPPs (ACPPs).....	12
Figure 1. 5 Evolution of a PC-3 Prostate Cancer Cell-Selective Cell Penetrating Peptide	14
Figure 1. 6 Protein Resurfacing Strategies and Cellular Uptake	16
Figure 1. 7 Identification of an Endosomolytic Peptide for Increased Delivery of Supercharged Proteins to the Cytosol of Mammalian Cells	19
Figure 1. 8 Cellular Delivery by Toxins	20
Figure 1. 9 Vault Nanoparticles for Protein Encapsulation	25
Figure 1. 10 Engineered Phage as Spatially Separated Nanocarriers	26
Figure 2. 1 Identification of Ypep.....	42
Figure 2. 2 Ypep ‘hot spot’ identification and optimization.....	47
Figure 2. 3 Live cell fluorescence microscopy images of PC-3	48
Figure 2. 4 Flow cytometry data comparisons of Ypep with penetratin and Tat.....	49
Figure 2. 5 Flow cytometry data showing the amount of internalized GFP in PC-3 cells or Hek-293 cells	51
Figure 2. 6 Efficiency of nLuc delivery to human prostate cancer cells (PC-3)	52
Figure 3. 1 Typical protein structure for nanobodies.....	61
Figure 3. 2 Resurfacing and purification of polycationic nanobodies	63

Figure 3. 3 Polycationic Resurfaced Nanobodies Reach the Cytosol of Mammalian Cells.....	65
Figure 3. 4 Polycationic resurfaced nanobodies retain function and stability	68
Figure 4. 1 Typical Antibody Structure.....	77
Figure 4. 2 Cartoon depiction of IgG and fragments discussed in this Chapter.	80
Figure 4. 3 Utility of scFv fusions in SunTag Technology.....	82
Figure 4. 4 The Architecture of hIgG and Nanobodies.....	85
Figure 4.5 Recent Applications of Nanobodies	88
Figure 4.6 Structural comparison of IgG and Monobodies.	90
Figure 5. 1 Crystal structure of BC2 nanobody interacting with its BC2T epitope.....	106
Figure 5. 2 Comparisons of Antibodies and Nanobodies	108
Figure 5. 3 Analysis of Nanobodies in ELISAs.....	111
Figure 5. 4 Evaluation of Nanobodies for use in Flow Cytometry Experiments.....	113
Figure 5. 5 Evaluation of Nanobodies for Western Blot Applications	115
Figure 6. 1 Overview of Intracellular NanoBit Reassembly.....	122
Figure 6. 2 Overview of <i>In Vivo</i> Intracellular NanoBit Assay	127
Figure 6. 3 <i>In Vitro</i> Mixing of Purified Proteins Produces Luminescence.....	129
Figure 6. 4 <i>In Vivo</i> Intracellular NanoBit Assay.....	130
Figure 6. 5 Amplification Tags for Imaging.....	135
Figure 6. 6 Imaging Translation Dynamics in Living Cells	135
Figure 6. 7 Evolution of U1A protein from U1hpII recognition to recognition of TAR RNA ..	139
Figure 6. 8 TAR RNA hairpins used to make 7X TAR repeat.....	140
Figure 6. 9 Image of Bead Loaded sfGFP-TBP6.7.....	141
Figure 6. 10 Use of TBP6.7 in Cells Generating TAR RNA Construct	142

Figure 6. 11 HIV-1 Helix Bundle for Live Cell Imaging	143
Figure 6. 12 <i>In Vivo</i> Cell Imaging with 5-helix/C-Peptide System	144
Figure S2. 1 Representative flow cytometry data from Figure 2.2.....	160
Figure S2. 2 MTT cell viability assay data	161
Figure S2. 3 Fluorescence microscopy comparison of Ypep	162
Figure S2. 4 Penetration of G4N-GFP at 37 °C or 4 °C	163
Figure S2. 5 Representative flow cytometry data from Figure 2.4.....	164
Figure S2. 6 Representative flow cytometry data from Figure 2.5.....	165
Figure S2. 7 PAGE analysis of all proteins used in this work.....	166
Figure S3. 1 SDS-PAGE of Purified Proteins.....	174
Figure S3. 2 Brightfield Images of Mammalian Cells.....	175
Figure S3. 3 Fluorescent Microscopy Images of Supercharged GFP Variants	176
Figure S3. 4 ITC Data.....	177
Figure S5. 1 ELISA-HRP.....	188
Figure S5. 2 ELISA – NanoLuciferase	189
Figure S5. 3 Representative Flow Cytometry Histograms.	192
Figure S5. 4 Selectivity for Epitope Validated <i>via</i> Western blot.....	193
Figure S5. 5 SDS-PAGE of BC2 nanobodies utilized in this work.....	194
Figure S6. 1 SDS-PAGE of Purified Proteins – NanoBit.....	203
Figure S6. 2 Live-Cell Fluorescent Microscopy Images – NanoBit.....	204
Figure S6. 3 SDS-PAGE of Proteins Used in this Work – Imaging Constructs.....	209

CHAPTER ONE

Inside Job: Methods for Delivering Exogenous Proteins to Mammalian Cells¹

1.1 Introduction

Currently, 7 of the top 10 selling drugs are biologics – all proteins. Their large size, structural complexity, and molecular diversity often results in surfaces capable of potent and selective recognition of receptors that challenge, or evade, traditional small-molecules. However, most proteins do not penetrate the lipid bilayer exterior of mammalian cells. This severe limitation dramatically limits the number of disease-relevant receptors that proteins can target and modulate. Given the major role proteins play in modern medicine, and the magnitude of this limitation, it is unsurprising that so much effort has been dedicated to overcoming this pesky impediment. In this first chapter, we summarize and evaluate current approaches for intracellular delivery of exogenous proteins to mammalian cells, and in doing so, aim to illuminate fertile ground for future discovery in this critical area of research and the potential impact of my graduate work.

It is difficult to overstate the importance of intravenously or subcutaneously administered proteins play in improving human health, and fighting off disease. Over the course of the last two decades, we have been witness to rapid growth in the biologics sector (principally proteins (antibodies)) of the pharmaceutical industry.¹⁻³ This is a very recent trend. For over a century, small-molecules were virtually the only game in town. Why have proteins emerged as major players in the pharmaceutical industry, as well as basic research? The answer to this question is

¹ Adapted from: Bruce, V.J. and McNaughton, B.R., *Cell Chemical Biology*, **submitted 2017**

simple: absolute necessity. The problem with small-molecules is that they are small and often require a hydrophobic binding pocket for recognition.⁴ While small-molecules will likely *always* play a prominent role in pharmaceutical science, most protein receptors lack well-defined small-molecule binding sites, and thus challenge or evade small-molecule recognition. In contrast, the relatively large size, structural diversity, and molecular complexity of proteins often generates surfaces capable of recognizing disease-relevant receptors that defy their small-molecule counterparts.⁵ In addition, the evolution of proteins that bind a targeted receptor, either in the laboratory or *in vivo*, is generally higher-throughput and simpler than the analogous small molecule discovery process. Finally, relatively recent advances in large scale protein expression and purification have made possible the production of protein drugs on pharmaceutically-relevant scales. However, it is widely known that most proteins do not appreciably penetrate the lipid bilayer exterior of mammalian cells. As a result, exogenously administered protein therapeutics are generally limited to acting on receptors that reside on the surface of mammalian cells.

Small-molecule drugs that act on an intracellular receptor must be hydrophilic enough to dissolve in water – since we are mostly water – but hydrophobic enough to pass through the lipid-bilayer membrane of mammalian cells.⁶⁻⁷ Practitioners of medicinal chemistry are familiar with the term ‘partition coefficient’ (logP), the distribution of a small-molecule in a mixture of water and octanol. This experiment essentially says something about the probability of a small-molecule to traverse the lipid bilayer membrane. Medicinal chemists can tinker with the atomic composition of a drug lead until cell penetration is optimized. However, analogous manipulation of proteins is daunting territory. Protein folding is largely driven by the collapse of hydrophobic residues to the interior, and concomitant display of hydrophilic residues on the surface. Dramatic resurfacing of proteins with hydrophobic amino acids, in an effort to mimic medicinal chemistry efforts in the

small-molecule world, often results in decreased protein stability and/or aggregation. As a result, researchers have been forced to develop conceptually new approaches for the delivery of exogenous proteins to the interior of mammalian cells. Here, we present several of these approaches, discuss their existing challenges and hopefully offer insight into the future of these research and discovery efforts.

1.2 Taking a Step Back: Delivering Nucleic Acids to Mammalian Cells

DNA, and mRNA, encode the amino acid composition of proteins. If either of these reagents can be delivered to the interior of a mammalian cell, it can be acted upon by endogenous transcription/translation machinery, ultimately resulting in the synthesis of a new intracellular protein. However, the polyanionic nature of nucleic acids prohibits direct passage across lipid bilayers. Common methods for intracellular nucleic acid delivery include: cationic lipids and polymers, viral carriers, and lipid bilayer deformation / physical insertion (**Figure 1.1A-C**).

1.2.1 Cationic Lipids and Polymers

Lipid cation transfection reagents are synthetic carriers that typically consist of a quaternary amine cation and lipid tail – usually alkyl chains or hydrophobic molecules like cholesterol.⁸ In aqueous solutions, these reagents spontaneously align into bilayer sheets with buried lipids and exposed cations. When mixed with DNA, ion-paired assembly between negatively charged phosphates on DNA and the cationic surface of the bilayer sheet enables capture and encapsulation of DNA (**Figure 1.1A**). DNA containing lipid vesicles can permeate lipid bilayers of mammalian cells, and essentially serve as a Trojan horse delivery reagent for DNA. The basic mechanism of lipid-cation transfection of DNA consists of: (A) formation of DNA-encapsulated lipid cation

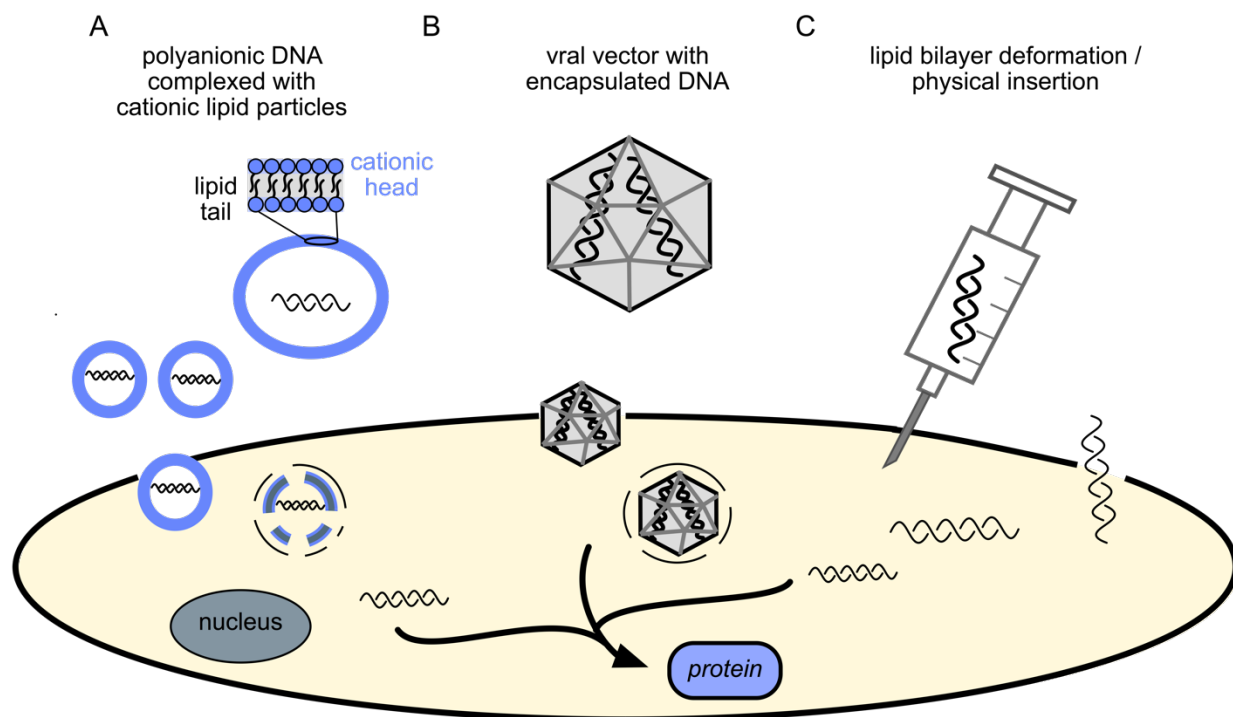


Figure 1. 1 Common Methods for Achieving Intracellular Delivery of Protein-Encoding Nucleic Acids. A) DNA/lipid-cation assemblies; B) Viral carriers with encapsulated DNA, C) Lipid bilayer deformation / physical insertion.

vesicles; (B) endocytosis of the lipid cation-DNA vesicles; (C) release of encapsulated DNA to the cytosol; (D) entry into the nucleus, followed by transcription and translation using endogenous machinery (**Figure 1.2**).

The first example of lipid-cation mediated delivery of DNA to mammalian cells was only reported in 1980.⁹⁻¹⁰ Since then, a multitude of lipid cation transfection reagents have become available commercially, and this method remains popular in mammalian cell culture.¹¹⁻¹² While these reagents protect nucleic acids from endogenous nucleases, and facilitate cell uptake, perhaps the greatest hurdles for broad use *in vivo* include: efficient escape from endosomes and innate toxicity. Lipid-cation mediated uptake often proceeds via endocytosis. Thus, efficient delivery of DNA requires subsequent escape from the endosomes to the cytosol (and then nucleus).¹³

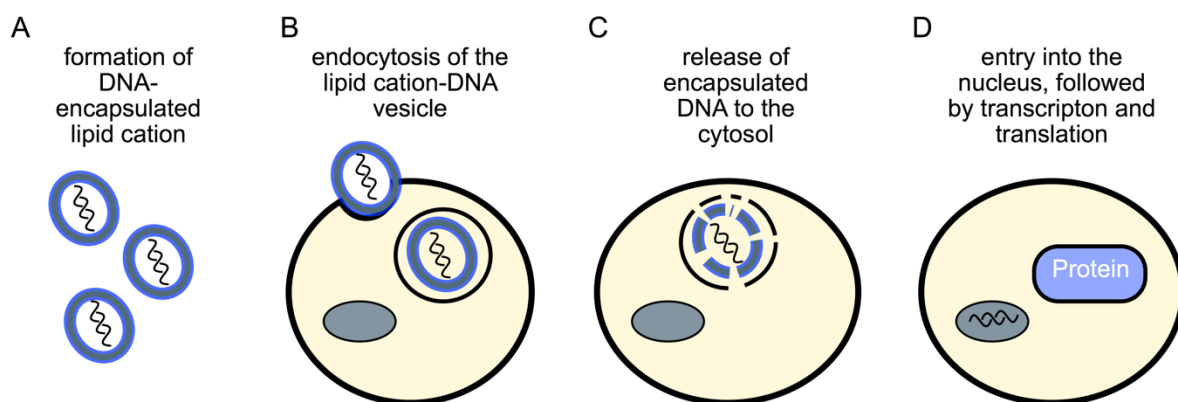


Figure 1. 2 Uptake Mechanism for Nucleic Acid Delivery with Cationic Lipids. A) First, polyanionic DNA is complexed with cationic lipids to form vesicles. B) External cations of the vesicles interact with the negatively charged membrane to induce endocytosis. C) DNA is released from the cationic vesicle and endosomes to reach the cytoplasm. D) From the cytoplasm, the DNA must enter the nucleus where it can be transcribed and translated into the desired protein.

Researchers have recently shown that controlling vesicle composition, by virtue of mixtures of lipid cation reagents, can increase the delivery of nucleic acids that end up in the cytosol.¹⁴ As endosomes mature into lysosomes, the internal pH drops from ~ 6.5 to ~ 4.5 .¹⁵⁻¹⁶ This change has been exploited as an environmental cue to enhance endosomal escape. For example, ionizable lipid nanocarriers are neutral at physiological pH (pH ~ 7) but become cationic in the acidic environment associated with late-stage endosomes / lysosomes. The resulting change in charge leads to favorable disruption of the endosomal membrane, and release of internalized nucleic acid into the cytosol (**Figure 1.3**). In one recent example, a new series of ionizable lipids were derived from naturally occurring aminoglycoside tobramycin and a biodegradable diester linker.¹⁷ Alternatively, Mayworth, McKinlay and co-workers characterized so-called charge-altering releasable transporters (CARTs), which are predicted on the use of cellularly biodegradable materials for the delivery of mRNA into cells.¹⁸ Once endocytosed these lipids undergo degradation, releasing polyanionic nucleic acids that disrupt, and escape from, the endosome membrane.

When used *in vivo*, lipid cation transfection reagents can induce an unwanted immune response, and thus are appreciably toxic. To help mitigate these affects, biodegradable transporters have been reported. A recent example utilized poly (β -amino esters) (PBAE) for the delivery of DNA to the brain.¹⁹ Researchers reported that this rapidly biodegradable nanocarrier resulted in minimal cytotoxicity or inflammatory response, while providing highly efficient DNA delivery, and rapid diffusivity within the brain parenchyma.

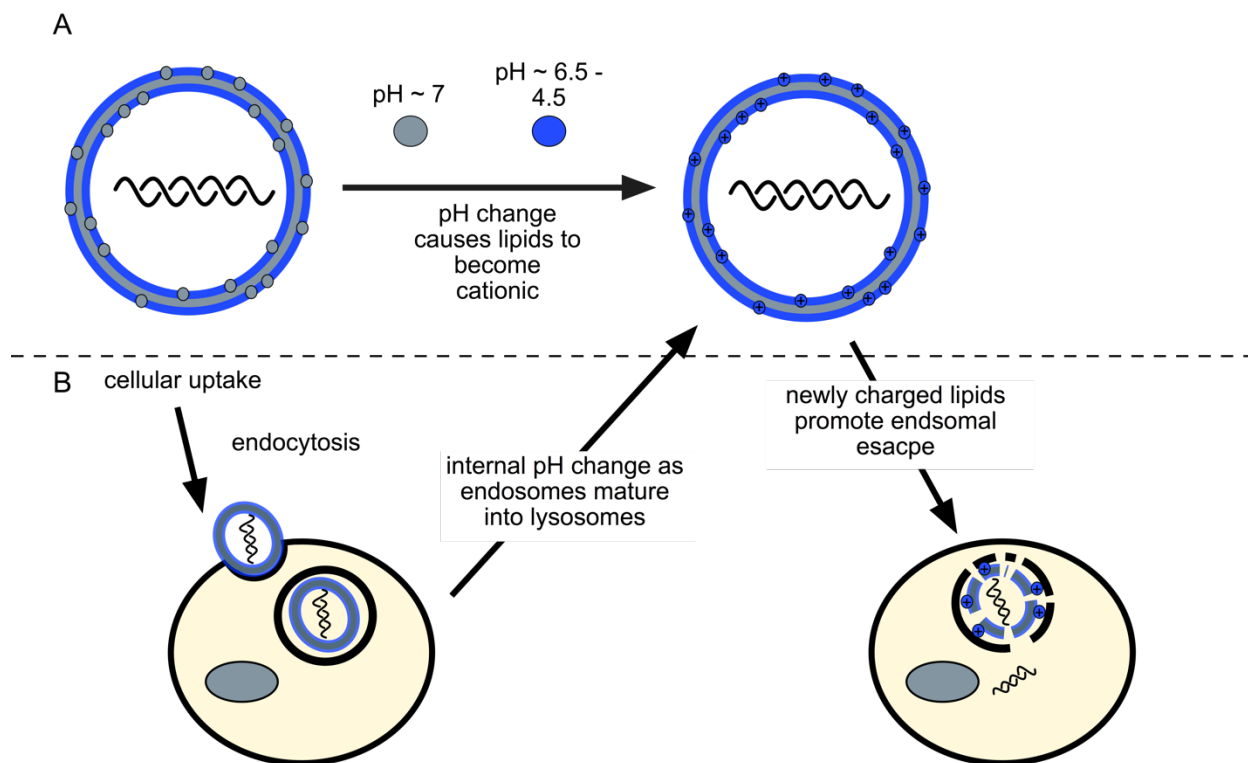


Figure 1. 3 Ionizable Lipid Nanocarriers. A) Ionizable lipids are mixed with typical cationic lipid reagents to form nanocarriers around DNA. A change in pH causes these once neutral lipids to become cationic. B) Cellular mechanism: Once endocytosed, the pH naturally becomes acidic within endosomal vesicles ionizing these lipids to their cationic form, thus promoting endosomal escape.

Despite recent advances, lipid-cations are principally used for the delivery of nucleic acids in mammalian cell culture. However, as researchers better understand the molecular dictates of

endosomolytic release, and overcome toxicity, these reagents might find expanded use for *in vivo* delivery of nucleic acids that encode therapeutically-relevant proteins.

1.2.2 Viral Carriers for the Intracellular Delivery of Protein-Encoding Nucleic Acids

In order to proliferate, viruses infect mammalian cells and transfer their genetic material. Unsurprisingly, perhaps, researchers have exploited Nature's solution of genetic transfer to mammalian cells as a method to deposit exogenous DNA across the lipid bilayer. Popular viral carriers include the retrovirus, lentivirus, adenovirus and adeno-associated virus (AAV).²⁰⁻²¹ Some vectors, such as AAV, can stably integrate DNA into the host genome at a specific site. While retroviruses can also integrate their DNA into its hosts genome, its occurrence is far more random, presenting the possibility of deleterious random insertion, which, if unlucky, could lead to diseases such as cancer.

Overall, the construction of viral carriers requires that the viral vector retains the necessary components for highly efficient genetic delivery, while at the same time being sufficiently mutated so that the pathogenic components have been silenced or replaced with therapeutic genes.²⁰ Viruses have had billions of years to evolve an ability to transfer genetic material to mammalian cells. As a result, they are highly efficient, and this approach has been the focus of a number of clinical trials. Leading clinical trials utilizing viral delivery include Hematopoietic-stem-cell gene delivery, liver directed gene therapy, T-cell immunotherapy for cancer and retinal gene therapy.²² Viral delivery of protein-encoding nucleic acids is particularly useful for diseases in which a malfunctioning, or missing, enzyme is replaced (generally referred to as enzyme replacement therapy). Of course, it is likely that an enzyme is only malfunctioning or completely missing in certain cells, and as a result, DNA encoding the proper protein should only be delivered to those

particular cells. Cell-selective delivery of nucleic acids is a focus of modern research, and advances need to be made in this area. Liver-selective delivery of gene-encoding nucleic acids can be achieved with viral carriers. One example is the treatment for severe hemophilia B, a bleeding disorder resulting from genetic mutation of a serine protease critical for blood clotting (factor IX or FIX).²³⁻²⁴ Researchers have shown that a single peripheral-vein infusion with a FIX-encoding AAV vector led to liver specific uptake, long-term expression of FIX (enzyme replacement) and a decreased risk of spontaneous hemorrhage.²⁵

While transfer of genetic information to mammalian cells, via viral carriers, is efficient, limitations and challenges exist. Foremost among these limitations and challenges are toxicity (often in the form of immunogenicity), transient versus permanent transfer of genes, restricted recognition of specific cell types, neutralization by endogenous antibodies, and restrictions on DNA carrying capacity. To help mitigate unwanted toxicity/immunogenicity, researchers have designed hybrid delivery vectors through the combination of retrovirus-like particles with synthetic polymers or lipids, essentially hiding the virus carrier from the immune system.

1.2.3 Lipid Bilayer Deformation / Physical Insertion

Physical methods that transiently disrupt the cellular membrane can be utilized for the delivery of nucleic acids to the interior of mammalian cells without the need of a chemical or viral carrier. Two of the most common methods include electroporation and microinjection.

Electroporation: Similar to lipid cation transfection, electroporation enjoys broad use as a method to deliver nucleic acids to the interior of cultured mammalian cells. Electroporation utilizes high-voltage electrical pulses to transiently disrupt the cell membrane. The resulting nanoscale pores in the plasma membrane allow relatively rapid uptake of exogenous materials, including

nucleic acids (and gene-encoding DNA). Perhaps unsurprisingly, electroporation is often toxic, and as a result, this approach to nucleic acid delivery is typically limited to mammalian cell culture.

However, electroporation has also found use as a means of delivery for nucleic acids in *ex vivo* applications, wherein isolated cells are genetically modified (by addition of exogenous nucleic acids), and selected for function (such as expression of a new protein), then reintroduced or implanted into the patient to simulate or restore damaged tissues or induce an immune response.²⁶ One of the greatest limitations for physical methods such as electroporation is low-throughput. The advent of microfluidic technologies and methodologies for increasing efficiency and magnitude of targeted cells has become a focus. Lu, Geng, and co-workers recently developed a protocol for continuous cellular transfection, reaching a liter volume of cells. The electric pulse is not applied to stagnant cells as in a classic electroporation cuvette, but to a flowing population resulting in increased transfection efficiency.²⁷

Microinjection: As the name implies, the process of microinjection requires direct physical insertion of DNA to the cytoplasm or nucleus of a mammalian cell via microneedles. The first microinjection of non-native DNA occurred in 1980 when glass micropipettes were used to introduce the Herpes simplex virus thymidine kinase gene to cultured mammalian cells.²⁸ Although this method is laborious, requires a skilled practitioner, and is very low-throughput; highly efficient delivery is achieved. Nanofabrication methods have enabled platforms that facilitate much higher-throughput microinjection of nucleic acids to mammalian cells.²⁹ For example, Shalek, Park, and co-workers utilized surfaced-modified vertical silicon ‘microneedle’ nanowires as a general platform for the introduction of nucleic acids and other biomolecules into immortalized and primary mammalian cells.³⁰

Microneedle-based approaches have been used to deliver small-molecule drugs, peptides, proteins, inactivated virus particles as well as plasmid DNA to the interior of mammalian cells.²⁹ Researchers have used various fabrication methods to evaluate a variety of ‘microneedles’ with different properties, including composition. Microneedles that are solid, hollow, dissolving, or made from hydrogel frameworks have been evaluated. This method is particularly useful for the delivery of plasmid DNA, along with other molecules and materials, into cell lines that are difficult to transfect with commercially available reagents such as cationic lipids. For example, researchers have used microneedles for the delivery of protein encoding plasmid DNA into primary neurons.³¹ For the delivery of nucleic acids via microneedles a review by Yuan, Chen and co-workers has been recently published.²⁹

Additionally, outside the realm of nucleic acids that encode proteins (DNA and mRNA), delivery of single-stranded antisense oligonucleotides (ASOs) and small interfering RNA (siRNA), both for silencing the production of particular proteins (which will not be discussed here but has been elsewhere), is an exploding trend that utilizes many of the same methods for delivery.³² Excellent reviews for their delivery can be found here (ASOs³³; siRNA³⁴).

1.3 Methods for Delivering Exogenous Proteins to Mammalian Cells

Conceptually, in comparison to the delivery of gene-encoding nucleic acids, it is simpler to *directly* deliver an exogenous protein to the interior of a mammalian cell. However, like nucleic acids, the size and theoretical net charge of proteins (most often with surfaced exposed charged residues) limits their cellular uptake.³⁵ Indeed, most naturally occurring proteins do not appreciably penetrate the lipid bilayer of mammalian cells. Concomitant with the rise in biologics as a major sector in the global pharmaceutical industry, researchers have dedicated an enormous

amount of effort to developing strategies for exogenous protein delivery to mammalian cells. Some of these methods are reviewed below.

1.3.1 Cell-Penetrating Peptides

Cell-penetrating peptides (CPPs) or protein transduction domains (PTDs) are relatively short peptides (<20 amino acids), which, when fused to a protein endow transport of that protein across the lipid bilayer membrane. First described in the 1980's, this method is probably one of the most straightforward platforms for direct protein delivery to the cell interior.³⁶

A selection of the most widely utilized CPPs include: Tat peptide from the transactivating protein Tat of HIV-1 (GRKKRRQRRRPPQ)³⁷⁻³⁸; the third helix of the homeodomain of *antennapedia* called penetratin (RQIKIWFQNRRMKWKK)³⁹; and polyarginine (R₉)⁴⁰. A quick glance at these sequences reveals a common theme: a high ratio of cationic residues (arginine and/or lysine). It has been known for some time that cationic residues interact favorably with the exterior of a mammalian cell membrane⁴¹, principally by engaging negatively charged glycosaminoglycans (GAGs) such as heparin sulfate (HS) and chondroitin sulfate (CS) on the surface of cells.⁴²⁻⁴⁵ Generally, arginine rich peptides are more potent carriers than their lysine rich counterparts, due to more favorable, bidentate interactions with sulfated proteoglycans on the cell surface.

1.3.2 Cell-Selective Cell-Penetrating Peptides

One limitation of polycationic CPPs is that they indiscriminately engage, and penetrate, mammalian cells (and drag in associated cargo). Ideally, CPPs would only penetrate, and deliver associated cargo, to diseased cells. One approach to achieving cell-selective delivery is cancer-

cell-Activated CPPs, (ACPPs), which have been reported by Tsien and co-workers (**Figure 1.4**).⁴⁶⁻

⁴⁷ These reagents consist of three components: (1) a polycationic CPP; (2) a polyanionic biopolymer that binds to, and essentially quenches the polycationic CPP, and; (3) a linker that can be cleaved by an excreted protease associated with diseased (commonly cancer) cells. Since this reagent is entirely composed of polypeptides, it can be genetically fused to a protein of interest. In the absence of cancer cell associated proteases (such as matrix metalloprotease 9, MMP-9) the polyanionic portion quenches the charge of the polycationic portion – disabling cell-penetration. However, in the presence of a cancer cell associated protease, the linker is cleaved, ultimately

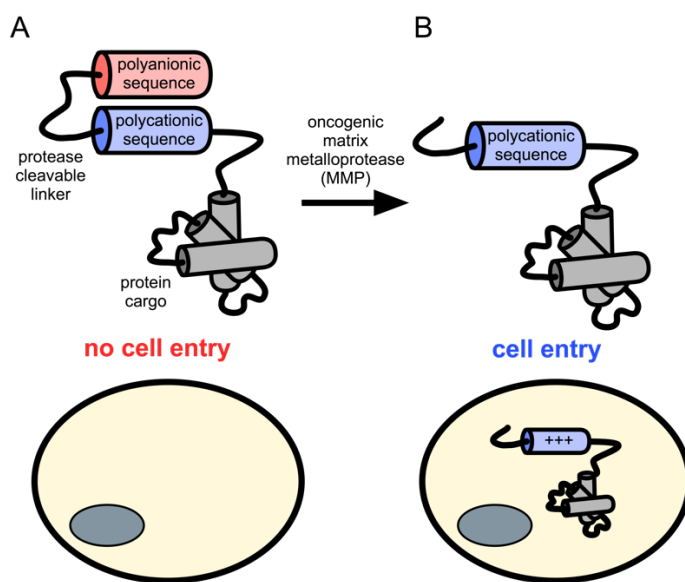


Figure 1. 4 Cancer Cell Activated CPPs (ACPPs). A) In the absence of cancer cell associated proteases, ACPPs remain inactive and unable to penetrate mammalian cells. B) Activation of ACPPs by the presence of cancer cell associated proteases (ex: MMP-9), cleaves the anionic component from the cationic CPP – protein fusion component, allowing the cationic CPP-protein fusion to penetrate neighboring (diseased) cells.

resulting in separation of polyanionic and polycationic-components. Following dissociation of the polyanionic component, the polycationic CPP-protein fusion can penetrate neighboring (diseased) cells.⁴⁸

An alternative strategy is to use laboratory evolution to select for cell-selective uptake. A number of labs⁴⁹⁻⁵¹, including our own⁵², have used rounds of positive and negative selection to evolve CPPs that selectively penetrate a cell line of interest. As an example, we performed positive/negative selections on a $\sim 10^{12}$ bacteriophage (phage) display 12mer-linear peptide library to identify phage (peptides) that potently penetrate PC-3 prostate cancer cells, but do not appreciably penetrate a number of non-prostate, or prostate but non-cancerous cell lines (**Figure 1.5**).⁵² Following multiple cycles of positive and negative selection, we identified Y-peptide, ‘Ypep’ (YTFGLKTSFNVQ), which selectively penetrates PC-3 cells.⁵³ Further efforts towards the optimization and characterization of this CPP will be more fully discussed in Chapter Two of this thesis.

1.3.3 Protein Resurfacing

As stated previously, the majority of reported CPPs contain a disproportionately high number of positively charged arginine and/or lysine residues, and have a high theoretical net charge. These positively charged residues facilitate interaction with negatively charged cell surface proteoglycans, such as sulfated proteoglycans – ultimately enabling cell uptake.⁵⁴ Protein engineers have extended the concept of polycationic-mediated delivery to the protein surface itself. By performing dramatic mutagenesis of protein surface residues to positively charged arginine and/or lysine, researchers have created varied degrees of polycationic resurfaced proteins, which contain a disproportionately high number of positively charged residues on the surface. These proteins penetrate mammalian cells, often with improved potency, compared to their polycationic CPP cousins.

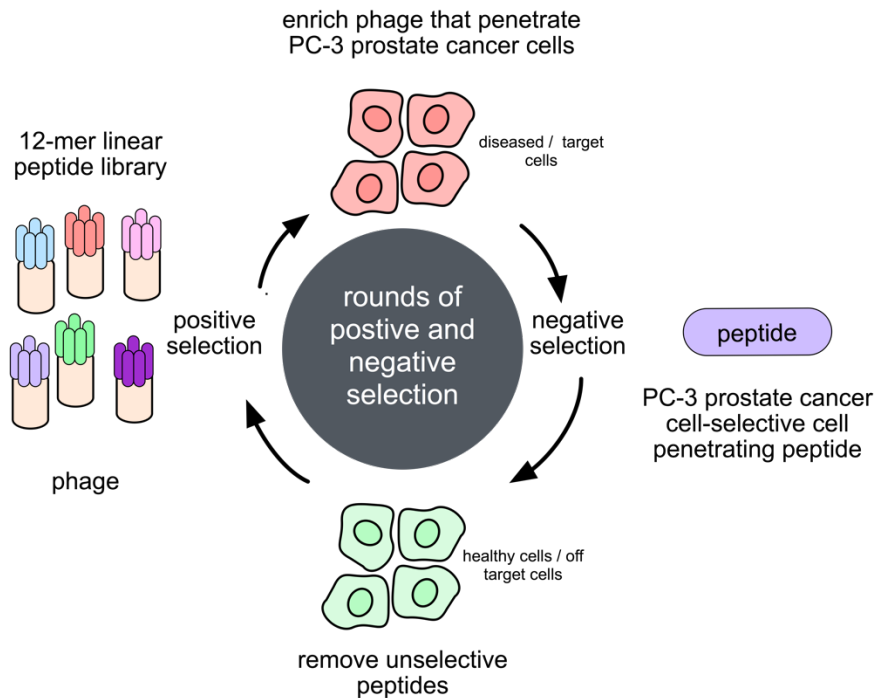


Figure 1.5 Evolution of a PC-3 Prostate Cancer Cell-Selective Cell Penetrating Peptide. In a positive selection, phage are incubated with PC-3 prostate cancer cells and enriched for cell penetration. Next, in a negative selection, cell-penetrating phage are incubated with off-target cells, and phage that don't penetrate are enriched.

In a seminal contribution, Raines and co-workers utilized a method they call 'arginine grafting' to turn proteins with no appreciable cell-penetration activity (RNase A or GFP) into potent cell-penetrating proteins.⁵⁵⁻⁵⁶ Arginine grafting is achieved by mutating clustered solvent exposed acidic residues to arginine (**Figure 1.6A** – GFP is shown as an example). The resulting arginine cluster can engage sulfated proteoglycans on the surface of mammalian cells – ultimately enabling cell uptake. In a therapeutically-relevant example, internalized RNase A led to cellular cytotoxicity, through destruction of intracellular RNA. However, like their polycationic CPP relatives, internalized proteins (GFP as an example) appear as punctate foci, suggesting a proportion reside within endosomes (**Figure 1.6B**).⁵⁷ Nonetheless, for highly efficient enzymatic

proteins (such as RNase A), a little goes a long way; endosome entrapment of a significant percentage of the population might not be a great concern.

In a conceptually related approach, Liu and co-workers pushed the limits of protein resurfacing by generating a series of ‘supercharged proteins’ – the result of extensive mutagenesis of surface residues to positively charged arginine and lysine (**Figure 1.6C**). In this case, the resulting proteins exemplified extreme stability, resistance to aggregation, and the ability to potently penetrate mammalian cells.⁵⁸ For example, researchers in the Liu laboratory ‘supercharged’ GFPs to have a theoretical net charge of +15, +25, or +36.⁵⁹⁻⁶⁰ Theoretical charge matters (at least in this context), and resulting uptake of supercharged GFP increases as a function of theoretical net charge. Subsequently, researchers in the Liu lab demonstrated that engineered ‘supercharged’ GFPs, and naturally occurring ‘supercharged’ proteins (theoretical net charge: molecular weight ratio of +1.27/kDa) can deliver genetically fused proteins to the interior of mammalian cells.⁶¹ Of great concern, is while these superpositively charged proteins are efficiently taken up by cells, the route of entry relies on endocytosis. As a result, an appreciable percentage of internalized (endocytosed) proteins remain sequestered in endosomes and cannot act on cytosolic or nuclear targets (**Figure 1.6D**).⁵⁷ However, researchers have recently developed methods to get endocytosed exogenous proteins outside of the vesicles, and into therapeutically relevant environments, including the cytosol and nucleus (See section 1.3.4) .

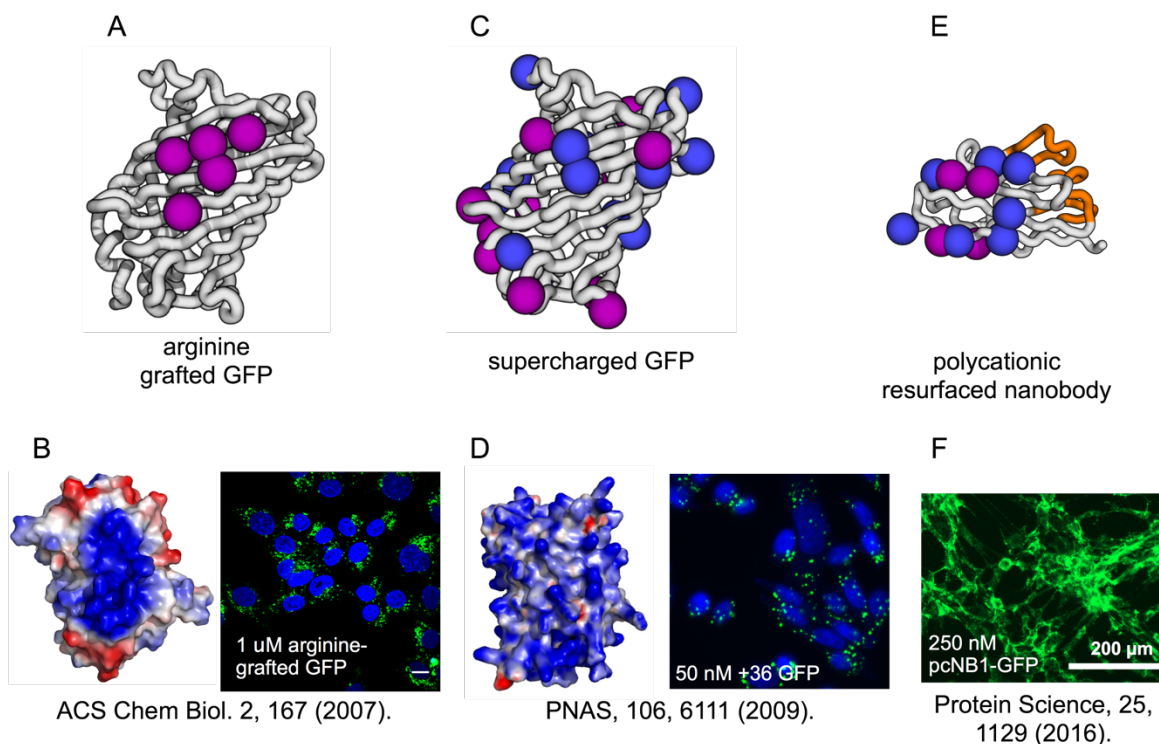


Figure 1. 6 Protein Resurfacing Strategies and Cellular Uptake. A) GFP shown as an example of arginine grafting, where a small solvent exposed patch has been mutated to arginines. B) Raines' arginine grafted GFP. Blue = cationic residues; red = anionic residues. Arginine grafted GFP penetrates mammalian cells, and largely appears as punctate foci, suggesting encapsulation within endosomes. C) Residues mutated to arginine (purple) and lysine (blue) to generate +36 GFP. D) Liu's supercharged +36 GFP. Blue = cationic; red = anionic. +36 GFP potentially penetrates mammalian cells, and largely appears as punctate foci, suggesting encapsulation within endosomes. In microscopy images, nuclei are stained blue with DAPI. E) Polycationic resurfacing of a nanobody (GFP binding nanobody shown as example). Residues mutated to arginine (purple) and lysine (blue) generate a theoretical net charge of +15. F) Fluorescence microscopy images of GFP fused nanobodies appear more evenly delivered, suggesting cytosolic delivery. (PDB: GFP: 1GFL; nanobody: 3OGO).

While polycationic protein resurfacing endows potent cell penetration, most proteins are not receptive to such extensive mutation. Additionally, CPPs are often limited in the amount of charge it can possess and still be beneficial for uptake, with polyarginine tags of greater than 20 residues decreasing uptake.⁶² Current work with 'supercharged' proteins suggests that polyarginines in the well-structured framework of a protein behave differently than polycationic CPPs.⁶³ Ideally, cell-penetrating drug leads and basic research tools could be prepared from a single polycationic

resurfaced protein architecture. Recently, we showed that nanobodies, the binding domain of camelid-derived heavy chain antibodies are amenable to cationic resurfacing.⁶⁴ Chapter Three of this thesis describes the characterization of these polycationic resurfaced nanobodies and their potential use as therapeutic drug leads.

Through a mechanism that is conceptually similar to nucleic acid delivery, Liu and co-workers recently showed that polyanionic resurfaced proteins (-30 GFP), when mixed with commercially available lipid cations, penetrate mammalian cells.⁶⁵ Concomitant with a rise of interest in delivering gene editing technologies, such as CRISPR, they show the co-complexation of purified Cas9 protein (+22 theoretical charge) with sgRNA yields an overall highly anionic supramolecular unit for protein/RNA encapsulation by cationic lipids and cellular uptake.

1.3.4 The Great Escape: Endosomolytic Peptides

Delivery of exogenous proteins to mammalian cells can sometimes occur via direct membrane translocation, which results in high levels of cytosolic residence of internalized cargo. Often, however, delivery occurs through one of the many well studied pathways for the uptake of exogenous material (including: endocytosis, micropinocytosis, and phagocytosis).⁶⁶⁻⁶⁷ Many protein delivery technologies proceed by an endocytotic pathway, and as a result, internalized cargo ends up trapped within endosomes. Functional delivery to the cytosol or nucleus then becomes a multi-step process: endocytosis followed by release of the cargo from endosomes into the cytosol (endosomolytic release).⁶⁸ In cases where nuclear delivery is required, fusion to a Nuclear Localization Sequence (NLS) is useful. Two of these three requirements (crossing the lipid bilayer and nuclear delivery, once in the cytosol) are, at least on a superficial level, solved

problems. What remains a significant challenge is deformation of the endosome and release of internalized cargo to the cytosol (or subsequently the nucleus, if necessary).

As stated previously, early stage endosomes mature into late stage endosomes, which then fully mature into lysosomes. Along this maturation, the pH inside each of these vesicles changes dramatically (from pH ~6.5 to pH ~4.5). Peptides – some naturally occurring and some evolved in the laboratory – can respond to this change, and in doing so, disrupt the internalized vesicle. This disruption ultimately enables escape of internalized cargo, such as the internalized protein. The peptide HA2 (GLFGAIAGFIENGWEGMIDGWYG) is one such example. It is derived from the hemagglutinin HA2 subunit of the influenza virus and is rich in glutamate residues.⁶⁹ In acidic environments associated with late stage endosomes/lysosomes, aspartate and glutamate residues become neutralized by protonation, and interact more efficiently with the endosomal membrane, resulting in deformation.⁷⁰⁻⁷¹ Alternatively, hydrophobic residues – specifically phenylalanine, in a peptide segment called the penetration-accelerating sequence (Pas; FFLIPKG) have been shown to increase cytosolic delivery of fused proteins.⁷²⁻⁷³ It has been suggested that addition of hydrophobic residues in Pas, in combination with CPPs (like polyarginine), enables uptake and endosomolytic activity, presumably through hydrophobic residues in Pas interacting with, and disrupting, endosomal membranes.

Recently, Liu, Li, and co-workers evaluated peptides classified under the umbrella of antimicrobial agents capable of penetrating microbial membranes, for an ability to enhance endosomal escape of exogenously delivered proteins (**Figure 1.7**).⁷⁴ From this study, researchers identified a 13-mer peptide (GLFDIIKKIAESF) called aurein 1.2 which enhances cytosolic delivery genetically fused proteins, both in live cell culture and *in vivo*. Interestingly, this peptide can selectively disrupt the endosomal membrane – facilitating escape of fused cargo protein from

the endosome to the cytosol – without disrupting the mammalian cell membrane or exhibiting appreciable toxicity. Since many protein delivery methods proceed via endocytosis, the pairing of non-toxic and highly efficient cell-penetrating and endosomolytic technologies will likely have a significant impact on the field of exogenous protein delivery.

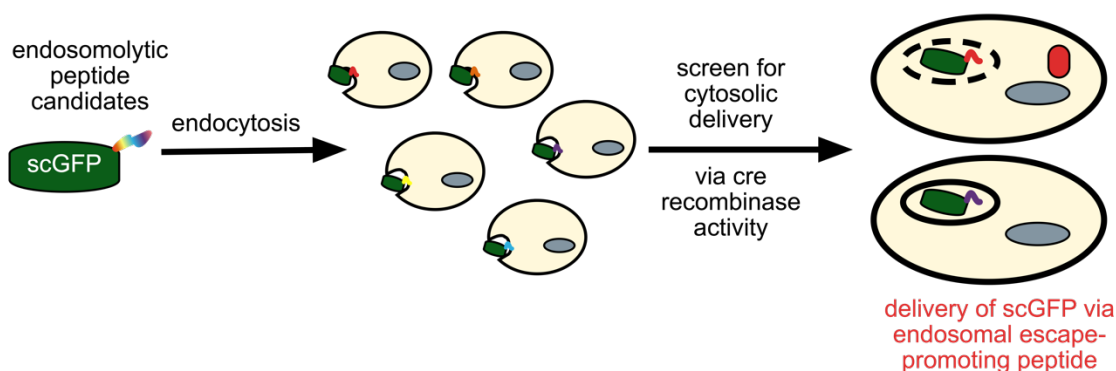


Figure 1. 7 Identification of an Endosomolytic Peptide for Increased Delivery of Supercharged Proteins to the Cytosol of Mammalian Cells. In this system, supercharged GFP was attached to cre recombinase, thus if cytosolic delivery was achieved, recombination of a reporter gene will trigger expression of a red fluorescent protein.

1.3.5 Employing Toxins to Deliver Exogenous Proteins

Protein assemblies from bacterial and plant toxins, such as anthrax, diphtheria and botulinum, potently penetrate mammalian cells and access the cytoplasm.⁷⁵ These toxins are of the type AB, and are comprised of two segments (**Figure 1.8A**). The A subunit typically refers to the cytotoxic element, which, once internalized, acts on its intracellular target. The B subunit, however, binds to specific cell surface receptors and enables uptake of the toxin into the cell – often through endocytosis as well as recognition by the A subunit. Once endocytosed, the B subunit interacts with the endosomal membrane. It is this interaction that results in poration of that membrane, and release of the toxin to the cytosol. Given that these toxins have been evolved to

penetrate the lipid bilayer of mammalian cells, *and*, escape endosomes, it is perhaps unsurprising that they are been exploited for exogenous protein delivery (**Figure 1.8B**).

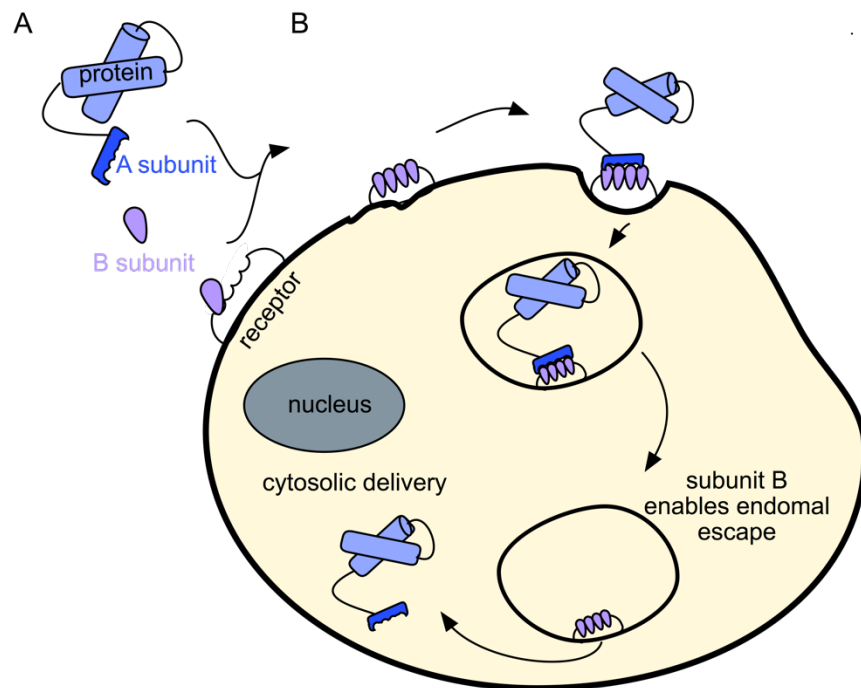


Figure 1. 8 Cellular Delivery by Toxins. A) AB type toxins are composed of an A subunit which typically delivers a toxin (or in this case protein) and a B subunit that recognizes a cell membrane receptor. B) Exogenous protein is delivered by endocytosis once the B subunit recognizes its cellular receptor. Next, subunit B generates pores to promote endosomal escape, thus delivering the exogenous protein to the cytosol.

Relatively recently, Botulinum neurotoxin type D (BoNT/D), inactivated by mutation of toxicity-inducing proteins, was exploited for the delivery of exogenous protein cargo to mammalian cells.⁷⁶ By genetic fusion of cargo proteins to the N-terminus of BoNT/D, researchers successfully delivered functional firefly luciferase and GFP to the cytoplasm of neuronal cells. However, simple genetic fusion of cargo protein to the terminus of inactivated toxin proteins can result in rather large proteins, and assemblies. For example, each BoNT/D fusion weighs approximately 150 kDa. A minimalistic approach is to remove the toxic component of the A

subunit leaving only what is required for recognition by subunit B, and simply work with a fusion between minimized subunit A and protein cargo. As a representative example, researchers have used the anthrax toxin as a bacterial transport nanomachine that can deliver proteins to mammalian cells. This system also belongs to the AB-toxin group and is made up of a protective antigen (PA, subunit B) component that recognizes the cell and an A subunit called the lethal factor protein (LF, 90 kDa) which is delivered to the cytosol of the cell through its recognition of PA. Recently the unnecessary cytotoxic components of subunit A (LF) were eliminated yielding a 30 kDa fragment (vs the wild type LF at 90 kDa) of the N-terminal domain (termed LF_N) that is still recognized by PA.⁷⁷ This system has been used to deliver antibody mimics as fusions between LF_N with the desired protein for delivery.⁷⁸ Recently, the Pentelute Laboratory commented on the functionality of this approach through the delivery of a monobody-LF_N fusion, and the directed inhibition of Bcr-Abl kinase activity and apoptosis in cultured cells.

Subunit B of bacterial and plant toxins initiates cellular internalization through selective recognition of cell surface receptors. These receptors include synaptotagmins enriched in neuronal cell surfaces for BoNT/D and tumor endothelial marker 8 or capillary morphogenesis gene 2 (TEM9 or CMG2, respectively) for PA.^{76, 79} While it can be beneficial under certain circumstances that one domain of the toxin exhibits selective binding to target tissue, it is also limiting as only certain tissues can be delivered to via this system. Recently, the PA/LF_N system has been modified to allow retargeting of the PA domain to receptors over expressed in tumor cells.⁸⁰ By fusing an engineered minimalistic high-affinity protein (Affibody⁸¹) specific for the HER2 receptor (a tyrosine kinase receptor over expressed in 20-25% of breast, gastric, and ovarian carcinomas) to the C terminus of a receptor recognition-deficient PA subunit (B subunit), targeted delivery of cytotoxic effectors was achieved to cell lines overexpressing HER2.⁸² Thus, such toxin platforms

might be highly tunable, as is the case with PA/LF_N, but prior knowledge concerning the system and cellular receptors available for delivery is required. Of great interest, is that these systems appear to efficiently escape the endosomes as a function of subunit B. Perhaps of concern for the general use of this system is its multi-component architecture. In order to achieve delivery, fairly large and multifaceted structures must be simultaneously supplied to target cells (PA: 83 kDa, LF_N: 30 kDa, + protein of interest).

1.3.6 Membrane – Disruption Based Techniques

Like electroporation and microinjection for the delivery of nucleic acids, membrane-disruption based techniques for the delivery of exogenous proteins act through transient disruption of the cell membrane. These methods take advantage of mechanical means, and include microneedles⁸³, electrical pulses⁸⁴, scrape/bead loading⁸⁵, acoustic sonoporation⁸⁶, chemical detergents and pore-forming agents⁸⁷⁻⁸⁸ and instrumental devices to make physical pores in the cellular membrane⁸⁹ that then allow the direct delivery of proteins to the cytoplasm. Physical methods are quite inconsistent with the biochemical assemblies previously discussed in this chapter. Considerations for endosomal escape and knowledge of appropriate receptors or surface interactions can be effectively eliminated because the delivery is direct and does not require carriers or chemical modification for protein delivery. Similar benefits and draw backs are experienced with these methods as with the previously discussed physical methods for nucleic acid delivery (benefits: temporal control over delivery dynamics, volume, and dosage concentration; limitations: low-throughput and limited use *in vivo*). Clinically speaking, these methods often result in a high proportion of cell death and thus must be accomplished *ex vivo*. Instrumentation and procedures based on increasing the flow and output for these methods are required for their

use in the clinic, and are perhaps becoming a reality with microfluidic devices. A very recent review by Jensen, Stewart and co-workers nicely discusses the current methods for membrane disruption techniques in terms of intracellular delivery and its outlooks.⁹⁰

The method of bead-loading is one example of these physical membrane-disruption based techniques. This technique utilizes glass beads (75 – 500 μM diameter) in a solution containing the desired macromolecule to be loaded into cells.⁸⁵ The solution of glass beads and protein are tapped onto the cells, creating transient pores in the cell membrane allowing macromolecules in the solution to be taken up by the cells. The glass beads, however, are too large for uptake and can be washed away from the cells. Recently, this technique was used deliver antibody Fab fragments for the imaging and detection of translation in living cells.⁹¹ A major advantage of this approach is that the loaded cells remain adherent and well-spread while endowing delivery of a large number of macromolecules to the cytosol of cells. However, while this technique is ideal for imaging cultured cells, it is not applicable for *in vivo* applications or methods requiring high-throughput.

The functional delivery of highly transient, nonreplicable protein agents as opposed to DNA sequences has many benefits and thus may offer improved specificity, increased safety and broader applicability. Through the delivery of folded proteins, control over therapeutic concentration and exposure time can be more precisely accomplished, as especially desired for many non-genetic disorders. Many of these methods must still overcome issues with cell selectivity and requirements concerning payload to become a validated means for the delivery of non-native therapeutic proteins, however the benefits afforded through the direct delivery of non-native protein therapeutics will no-doubt continue the push for work in this area.

1.4 Alternative Carriers for Protein Delivery

Carrier methods for the delivery of proteins can broadly be defined as any of the methods discussed here, but for the purpose of this chapter are physical entities that either encapsulate or act as a solid frame to protect desired proteins and include examples from delivery platforms such as protein nanocages and phage as delivery vehicles. Protein nanocages, such as vault proteins have been heavily studied by the Wender Laboratory. They are composed of protein subunits that can self-assemble to form a hollow interior to create a protein complex (**Figure 1.9**).⁹² It is this hollow interior that can encapsulate hundreds to thousands of proteins. Because they are made of amino acids and can be recombinantly expressed, studies have focused on modifying specific residues for precise encapsulation of desired cargo as well as external features to improve cellular specificity. A recent paper from the Wender Laboratory utilized lysine and cysteine residues to modify the exterior of these vault complexes with specific cell targeting peptides and imaging reagents for optimized cellular recognition and *in vitro* studies.⁹³ A conceptually similar procedure has been reported by the Hilvert Laboratory. They have engineered a capsid-forming enzyme lumazine synthase (AaLS-13) to create an altered electrostatic environment through the introduction of glutamate residues on the luminal surface giving the protein nanocarrier a highly negative charge.⁹⁴⁻⁹⁵ Upon mixing with proteins displaying a highly cationic charge (such as +36 GFP), efficient encapsulation of the polycationic GFP was achieved by the AaLS-13 carrier. Recently, researchers in the Hilvert Laboratory have shown that these nanocarriers can organize into multiple shells with variable packaging densities per each enclosure.⁹⁶⁻⁹⁷ We believe further work in this area will include the investigation of the use of these platforms for targeted therapeutic delivery applications. Another encapsulation method for the delivery of proteins to mammalian cells has been developed by the Rotello Laboratory and employs the use of nanoparticle-stabilized

capsules (NPSCs). Specifically, this method utilizes gold nanoparticles (AuNPs) in the presence of desired protein for delivery and a fatty acid capsule interior. This results in a stabilized container based on co-assembled supramolecular interactions between the components.⁹⁸ Recently, this scheme was used for the delivery of nanoassemblies consisting of Cas9 protein and sgRNA to cells for CRISPR gene-editing.⁹⁹ To achieve delivery the Cas9 protein was rationally engineered to contain an N-terminal poly-glutamic acid chain (1-20 residues) for association with arginine coated AuNPs and a C-terminal nuclear localization sequence (NLS) for nuclear delivery. They show efficient delivery of these nanoassemblies to HeLa cells for functional genome editing.

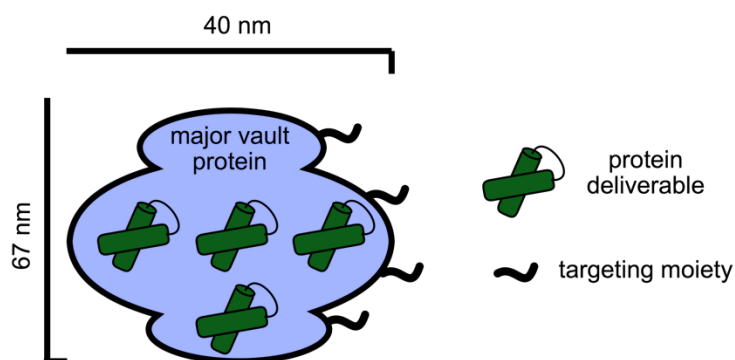


Figure 1. 9 Vault Nanoparticles for Protein Encapsulation. Nanoparticles are composed of 78 copies of the Major Vault Protein (MVP) and assembles into a 67 nm X 40 nm complex. Vaults can encapsulate and protect a variety of protein cargos as well as be engineered to display targeting moieties and cell-penetrating peptides.

Phage are an example of a specific bacterial virus that can be engineered to become a genetically defined nanocarrier. Of particular interest for this chapter are filamentous M13 phage which display genetically defined coat proteins.¹⁰⁰ This type of phage display one major coat protein called p8 and four minor coat proteins called p3, p6, p7, and p9, each rendering multiple copies of itself per a phage in a spatially defined manner. Through genetic and enzymatic manipulation of those displayed proteins our laboratory was able to generate nanocarriers for the intracellular delivery of functional exogenous protein to specific human prostate cancer cells (PC-

3).⁵² Our strategy relied on genetic modification of p3 and p9, such that p3 displayed a potent PC-3 CPP discussed above, and p9 displayed a biotin acceptor peptide (**Figure 1.10**).¹⁰¹ The acceptor peptide on p9 allowed us to accomplish *in vitro* biotinylation and further assemble via enzymatic conjugation any streptavidin fusion protein of imaging or therapeutic use to p9. As a proof-of-concept study, we show that this phage nanocarrier system can cell-selectively deliver a functional enzyme (horseradish peroxidase, HRP) to the interior of cells. When paired with an HRP substrate and the pro-drug 3-indol-acetic acid (IAA), appreciable cell death results, demonstrating the therapeutic utility of phage assisted delivery of HRP.

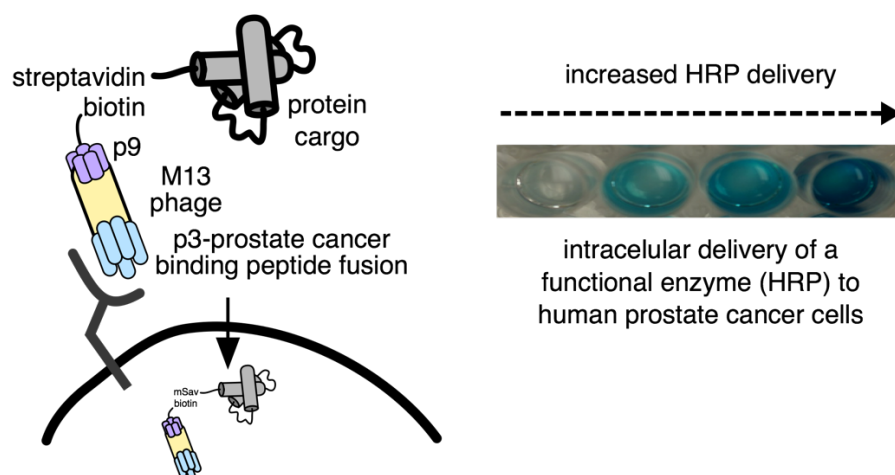


Figure 1. 10 Engineered Phage as Spatially Separated Nanocarriers. M13 phage can be genetically engineered to display a targeting peptide on p3 (CPP for PC3 cells in this example) and a biotin acceptor peptide on p9. Biotinylation of the acceptor peptide on p9 allows for conjugation of a streptavidin fusion protein (HRP in this case) to p9. Incubation of PC-3 cells with this phage functionalized nanocarrier and HRP substrate (TMB) results in a colorimetric reaction with increased color indicating increased delivery. This figure was adapted from *Bioconjugate Chem*, 2014, 25, 1620.

The carrier methods described here yield a well-defined structure on which to deliver proteins to mammalian cells. A highlight of these methods is their multifaceted nature for true versatility in delivery. Their inherent nature of spatial differentiation within one subunit endow them with a multitude of possible functionalization outcomes. Designing them with optimized features for

cellular uptake, precise delivery, and efficient payloads while minimizing cytotoxicity and immunogenicity are of greatest concern.

1.5 Conclusion

Proteins are an extraordinarily valuable reagent, whether its use be clinical or for basic research. The methods presented here are of significant pharmaceutical interest because they are laying the framework for the transport of highly desired biologic therapeutics into cells that cannot in their own right afford effective delivery. Cells, by virtue of their outer membrane, have become adept at preventing untargeted or non-natural complexes from entering. The unique dynamics found in the cellular membrane has rightly been found of interest for many academic groups as well as research and pharmaceutical companies. Concomitantly, the ever-increasing discovery and characterization of intracellular disease targets and tools, such as CRISPR, to regulate or treat disease states will continue to fuel aspirations towards systems for the delivery of exogenous proteins that meet the standards of clinical efficacy and bioavailability.

In this chapter, I have discussed methods for the indirect and direct delivery of exogenous proteins to mammalian cells and how each approach is in the process of overcoming limitations for increased efficiency and robustness. The first part of this thesis (Chapters Two and Three) highlight effort I have dedicated towards the obstacle of exogenous protein delivery through the optimization and characterization of a cell-penetrating peptide specific for prostate cancer cells and the use of polycationic resurfaced nanobodies as a potentially general scaffold for intracellularly targeted protein delivery. Finally, technologies for the analysis of exogenous protein delivery will be discussed in Chapter Six. This includes the design of a split nanoluciferase (NanoBit) assay for high-throughput characterization of the cytosolic delivery of proteins. We

believe this platform will allow for high-throughput screening and the identification of fundamentally novel methods for future exogenous protein delivery efforts. Moreover, advances within these techniques will be (and in some cases, are being) used to deliver gene-editing tools, including CRISPR, which requires the delivery of a two-part system including nucleic acid and protein reagents. As more tools and biologics are discovered that can greatly impact and treat human disease states, the methods for delivering these lifesaving reagents will play a prominent role in their use and development.

REFERENCES

1. Weiner, G. J., Building better monoclonal antibody-based therapeutics. *Nat Rev Cancer* **2015**, *15* (6), 361-70.
2. Evans, I., Follow-on biologics: a new play for big pharma: Healthcare 2010. *Yale J Biol Med* **2010**, *83* (2), 97-100.
3. Walsh, G., Biopharmaceutical benchmarks 2014. *Nat Biotechnol* **2014**, *32* (10), 992-1000.
4. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L., How many drug targets are there? *Nat Rev Drug Discov* **2006**, *5* (12), 993-6.
5. Chapman, A. M.; McNaughton, B. R., Scratching the Surface: Resurfacing Proteins to Endow New Properties and Function. *Cell Chem Biol* **2016**, *23* (5), 543-53.
6. Gupta, B.; Levchenko, T. S.; Torchilin, V. P., Intracellular delivery of large molecules and small particles by cell-penetrating proteins and peptides. *Adv Drug Deliv Rev* **2005**, *57* (4), 637-51.
7. Yang, N. J.; Hinner, M. J., Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol* **2015**, *1266*, 29-53.
8. Lonez, C.; Vandenbranden, M.; Ruyschaert, J. M., Cationic liposomal lipids: from gene carriers to cell signaling. *Prog Lipid Res* **2008**, *47* (5), 340-7.
9. Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M., Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A* **1987**, *84* (21), 7413-7.
10. Fraley, R.; Subramani, S.; Berg, P.; Papahadjopoulos, D., Introduction of liposome-encapsulated SV40 DNA into cells. *J Biol Chem* **1980**, *255* (21), 10431-5.

11. Juliano, R. L., The delivery of therapeutic oligonucleotides. *Nucleic Acids Res* **2016**, *44* (14), 6518-6548.
12. Akinc, A.; Zumbuehl, A.; Goldberg, M.; Leshchiner, E. S.; Busini, V.; Hossain, N.; Bacallado, S. A.; Nguyen, D. N.; Fuller, J.; Alvarez, R.; Borodovsky, A.; Borland, T.; Constien, R.; de Fougerolles, A.; Dorkin, J. R.; Jayaprakash, K. N.; Jayaraman, M.; John, M.; Koteliansky, V.; Manoharan, M.; Nechev, L.; Qin, J.; Racie, T.; Raitcheva, D.; Rajeev, K. G.; Sah, D. W. Y.; Soutschek, J.; Toudjarska, I.; Vornlocher, H. P.; Zimmermann, T. S.; Langer, R.; Anderson, D. G., A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nature Biotechnology* **2008**, *26* (5), 561-569.
13. Juliano, R. L.; Carver, K., Cellular uptake and intracellular trafficking of oligonucleotides. *Adv Drug Deliv Rev* **2015**, *87*, 35-45.
14. Colombani, T.; Peuziat, P.; Dallet, L.; Haudebourg, T.; Mevel, M.; Berchel, M.; Lambert, O.; Habrant, D.; Pitard, B., Self-assembling complexes between binary mixtures of lipids with different linkers and nucleic acids promote universal mRNA, DNA and siRNA delivery. *J Control Release* **2017**, *249*, 131-142.
15. Schmid, S.; Fuchs, R.; Kielian, M.; Helenius, A.; Mellman, I., Acidification of endosome subpopulations in wild-type Chinese hamster ovary cells and temperature-sensitive acidification-defective mutants. *J Cell Biol* **1989**, *108* (4), 1291-300.
16. Serresi, M.; Bizzarri, R.; Cardarelli, F.; Beltram, F., Real-time measurement of endosomal acidification by a novel genetically encoded biosensor. *Analytical and Bioanalytical Chemistry* **2009**, *393* (4), 1123-1133.
17. Habrant, D.; Peuziat, P.; Colombani, T.; Dallet, L.; Gehin, J.; Goudeau, E.; Evrard, B.; Lambert, O.; Haudebourg, T.; Pitard, B., Design of Ionizable Lipids To Overcome the Limiting

Step of Endosomal Escape: Application in the Intracellular Delivery of mRNA, DNA, and siRNA. *J Med Chem* **2016**, *59* (7), 3046-62.

18. McKinlay, C. J.; Vargas, J. R.; Blake, T. R.; Hardy, J. W.; Kanada, M.; Contag, C. H.; Wender, P. A.; Waymouth, R. M., Charge-altering releasable transporters (CARTs) for the delivery and release of mRNA in living animals. *Proc Natl Acad Sci U S A* **2017**, *114* (4), E448-E456.

19. Mastorakos, P.; Song, E.; Zhang, C.; Berry, S.; Park, H. W.; Kim, Y. E.; Park, J. S.; Lee, S.; Suk, J. S.; Hanes, J., Biodegradable DNA Nanoparticles that Provide Widespread Gene Delivery in the Brain. *Small* **2016**, *12* (5), 678-85.

20. El-Aneed, A., An overview of current delivery systems in cancer gene therapy. *Journal of Controlled Release* **2004**, *94* (1), 1-14.

21. Witlox, M. A.; Lamfers, M. L.; Wuisman, P. I. J. M.; Curiel, D. T.; Siegal, G. P., Evolving gene therapy approaches for osteosarcoma using viral vectors: Review. *Bone* **2007**, *40* (4), 797-812.

22. Naldini, L., Gene therapy returns to centre stage. *Nature* **2015**, *526* (7573), 351-360.

23. High, K. H.; Nathwani, A.; Spencer, T.; Lillicrap, D., Current status of haemophilia gene therapy. *Haemophilia* **2014**, *20*, 43-49.

24. Mingozzi, F.; High, K. A., Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* **2011**, *12* (5), 341-55.

25. Nathwani, A. C.; Tuddenham, E. G.; Rangarajan, S.; Rosales, C.; McIntosh, J.; Linch, D. C.; Chowdary, P.; Riddell, A.; Pie, A. J.; Harrington, C.; O'Beirne, J.; Smith, K.; Pasi, J.; Glader, B.; Rustagi, P.; Ng, C. Y.; Kay, M. A.; Zhou, J.; Spence, Y.; Morton, C. L.; Allay, J.; Coleman, J.; Sleep, S.; Cunningham, J. M.; Srivastava, D.; Basner-Tschakarjan, E.; Mingozzi, F.; High, K.

- A.; Gray, J. T.; Reiss, U. M.; Nienhuis, A. W.; Davidoff, A. M., Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med* **2011**, *365* (25), 2357-65.
26. Hui, S. W.; Li, L. H., In vitro and ex vivo gene delivery to cells by electroporation. *Methods Mol Med* **2000**, *37*, 157-71.
27. Geng, T.; Zhan, Y. H.; Wang, J.; Lu, C., Transfection of cells using flow-through electroporation based on constant voltage. *Nature Protocols* **2011**, *6* (8), 1192-1208.
28. Capecchi, M. R., High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* **1980**, *22* (2 Pt 2), 479-88.
29. Chen, W.; Li, H.; Shi, D.; Liu, Z. G.; Yuan, W. E., Microneedles As a Delivery System for Gene Therapy. *Front Pharmacol* **2016**, *7*.
30. Shalek, A. K.; Robinson, J. T.; Karp, E. S.; Lee, J. S.; Ahn, D. R.; Yoon, M. H.; Sutton, A.; Jorgolli, M.; Gertner, R. S.; Gujral, T. S.; MacBeath, G.; Yang, E. G.; Park, H., Vertical silicon nanowires as a universal platform for delivering biomolecules into living cells. *Proc Natl Acad Sci U S A* **2010**, *107* (5), 1870-5.
31. Wang, Y.; Yang, Y.; Yan, L.; Kwok, S. Y.; Li, W.; Wang, Z.; Zhu, X.; Zhu, G.; Zhang, W.; Chen, X.; Shi, P., Poking cells for efficient vector-free intracellular delivery. *Nat Commun* **2014**, *5*, 4466.
32. Dowdy, S. F., Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnology* **2017**, *35* (3), 222-229.
33. Crooke, S. T.; Wang, S. Y.; Vickers, T. A.; Shen, W.; Liang, X. H., Cellular uptake and trafficking of antisense oligonucleotides. *Nature Biotechnology* **2017**, *35* (3), 230-237.
34. Wittrup, A.; Lieberman, J., Knocking down disease: a progress report on siRNA therapeutics. *Nat Rev Genet* **2015**, *16* (9), 543-52.

35. Marschall, A. L. J.; Zhang, C. C.; Frenzel, A.; Schirrmann, T.; Hust, M.; Perez, F.; Dubel, S., Delivery of antibodies to the cytosol Debunking the myths. *Mabs* **2014**, 6 (4), 943-956.
36. Bechara, C.; Sagan, S., Cell-penetrating peptides: 20 years later, where do we stand? *FEBS Lett* **2013**, 587 (12), 1693-702.
37. Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988**, 55 (6), 1189-93.
38. Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **1988**, 55 (6), 1179-88.
39. Derossi, D.; Calvet, S.; Trembleau, A.; Brunissen, A.; Chassaing, G.; Prochiantz, A., Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. *J Biol Chem* **1996**, 271 (30), 18188-93.
40. Fuchs, S. M.; Raines, R. T., Polyarginine as a multifunctional fusion tag. *Protein Sci* **2005**, 14 (6), 1538-44.
41. Ryser, H. J.; Hancock, R., Histones and basic polyamino acids stimulate the uptake of albumin by tumor cells in culture. *Science* **1965**, 150 (3695), 501-3.
42. Fuchs, S. M.; Raines, R. T., Pathway for polyarginine entry into mammalian cells. *Biochemistry* **2004**, 43 (9), 2438-44.
43. Fuchs, S. M.; Raines, R. T., Internalization of cationic peptides: the road less (or more?) traveled. *Cellular and Molecular Life Sciences* **2006**, 63 (16), 1819-1822.
44. Poon, G. M.; Gariepy, J., Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells. *Biochem Soc Trans* **2007**, 35 (Pt 4), 788-93.

45. Richard, J. P.; Melikov, K.; Brooks, H.; Prevot, P.; Lebleu, B.; Chernomordik, L. V., Cellular uptake of unconjugated TAT peptide involves clathrin-dependent endocytosis and heparan sulfate receptors. *Journal of Biological Chemistry* **2005**, *280* (15), 15300-15306.
46. Aguilera, T. A.; Olson, E. S.; Timmers, M. M.; Jiang, T.; Tsien, R. Y., Systemic in vivo distribution of activatable cell penetrating peptides is superior to that of cell penetrating peptides. *Integr Biol* **2009**, *1* (5-6), 371-381.
47. Olson, E. S.; Aguilera, T. A.; Jiang, T.; Ellies, L. G.; Nguyen, Q. T.; Wong, E. H.; Gross, L. A.; Tsien, R. Y., In vivo characterization of activatable cell penetrating peptides for targeting protease activity in cancer. *Integr Biol* **2009**, *1* (5-6), 382-393.
48. Crisp, J. L.; Savariar, E. N.; Glasgow, H. L.; Ellies, L. G.; Whitney, M. A.; Tsien, R. Y., Dual Targeting of Integrin $\alpha(v)\beta(3)$ and Matrix Metalloproteinase-2 for Optical Imaging of Tumors and Chemotherapeutic Delivery. *Molecular Cancer Therapeutics* **2014**, *13* (6), 1514-1525.
49. Dickerson, T. J.; Kaufmann, G. F.; Janda, K. D., Bacteriophage-mediated protein delivery into the central nervous system and its application in immunopharmacotherapy. *Expert Opin Biol Th* **2005**, *5* (6), 773-781.
50. Frenkel, D.; Solomon, B., Filamentous phage as vector-mediated antibody delivery to the brain. *P Natl Acad Sci USA* **2002**, *99* (8), 5675-5679.
51. Ivanenkov, V. V.; Felici, F.; Menon, A. G., Targeted delivery of multivalent phage display vectors into mammalian cells (vol 1448, pg 463, 1999). *Bba-Mol Cell Res* **1999**, *1451* (2-3), 364-364.

52. DePorter, S. M.; Lui, I.; Mohan, U.; McNaughton, B. R., A protein transduction domain with cell uptake and selectivity profiles that are controlled by multivalency effects. *Chem Biol* **2013**, *20* (3), 434-44.
53. DePorter, S. M.; Lui, I.; Bruce, V. J.; Gray, M. A.; Lopez-Islas, M.; McNaughton, B. R., Mutagenesis modulates the uptake efficiency, cell-selectivity, and functional enzyme delivery of a protein transduction domain. *Mol Biosyst* **2014**, *10* (1), 18-23.
54. van Meer, G.; Voelker, D. R.; Feigenson, G. W., Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **2008**, *9* (2), 112-24.
55. Fuchs, S. M.; Rutkoski, T. J.; Kung, V. M.; Groeschl, R. T.; Raines, R. T., Increasing the potency of a cytotoxin with an arginine graft. *Protein Eng Des Sel* **2007**, *20* (10), 505-9.
56. Fuchs, S. M.; Raines, R. T., Arginine grafting to endow cell permeability. *ACS Chem Biol* **2007**, *2* (3), 167-70.
57. Thompson, D. B.; Villasenor, R.; Dorr, B. M.; Zerial, M.; Liu, D. R., Cellular uptake mechanisms and endosomal trafficking of supercharged proteins. *Chem Biol* **2012**, *19* (7), 831-43.
58. Lawrence, M. S.; Phillips, K. J.; Liu, D. R., Supercharging proteins can impart unusual resilience. *J Am Chem Soc* **2007**, *129* (33), 10110-2.
59. McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R., Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc Natl Acad Sci U S A* **2009**, *106* (15), 6111-6.
60. Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* **2010**, *5* (8), 747-52.

61. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R., A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. *Chem Biol* **2011**, *18* (7), 833-8.
62. Mitchell, D. J.; Kim, D. T.; Steinman, L.; Fathman, C. G.; Rothbard, J. B., Polyarginine enters cells more efficiently than other polycationic homopolymers. *J Pept Res* **2000**, *56* (5), 318-25.
63. Appelbaum, J. S.; LaRoche, J. R.; Smith, B. A.; Balkin, D. M.; Holub, J. M.; Schepartz, A., Arginine topology controls escape of minimally cationic proteins from early endosomes to the cytoplasm. *Chem Biol* **2012**, *19* (7), 819-30.
64. Bruce, V. J.; Lopez-Islas, M.; McNaughton, B. R., Resurfaced cell-penetrating nanobodies: A potentially general scaffold for intracellularly targeted protein discovery. *Protein Sci* **2016**, *25* (6), 1129-37.
65. Zuris, J. A.; Thompson, D. B.; Shu, Y.; Guilinger, J. P.; Bessen, J. L.; Hu, J. H.; Maeder, M. L.; Joung, J. K.; Chen, Z. Y.; Liu, D. R., Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* **2015**, *33* (1), 73-80.
66. Duchardt, F.; Fotin-Mleczek, M.; Schwarz, H.; Fischer, R.; Brock, R., A comprehensive model for the cellular uptake of cationic cell-penetrating peptides. *Traffic* **2007**, *8* (7), 848-66.
67. Ziegler, A., Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans. *Adv Drug Deliver Rev* **2008**, *60* (4-5), 580-597.
68. Erazo-Oliveras, A.; Muthukrishnan, N.; Baker, R.; Wang, T. Y.; Pellois, J. P., Improving the endosomal escape of cell-penetrating peptides and their cargos: strategies and challenges. *Pharmaceuticals (Basel)* **2012**, *5* (11), 1177-209.

69. Wharton, S. A.; Martin, S. R.; Ruigrok, R. W.; Skehel, J. J.; Wiley, D. C., Membrane fusion by peptide analogues of influenza virus haemagglutinin. *J Gen Virol* **1988**, *69* (Pt 8), 1847-57.
70. Tamm, L. K.; Han, X., Viral fusion peptides: a tool set to disrupt and connect biological membranes. *Biosci Rep* **2000**, *20* (6), 501-18.
71. Turk, M. J.; Reddy, J. A.; Chmielewski, J. A.; Low, P. S., Characterization of a novel pH-sensitive peptide that enhances drug release from folate-targeted liposomes at endosomal pHs. *Biochim Biophys Acta* **2002**, *1559* (1), 56-68.
72. Takayama, K.; Hirose, H.; Tanaka, G.; Pujals, S.; Katayama, S.; Nakase, I.; Futaki, S., Effect of the attachment of a penetration accelerating sequence and the influence of hydrophobicity on octaarginine-mediated intracellular delivery. *Mol Pharm* **2012**, *9* (5), 1222-30.
73. Lonn, P.; Kacsinta, A. D.; Cui, X. S.; Hamil, A. S.; Kaulich, M.; Gogoi, K.; Dowdy, S. F., Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Sci Rep* **2016**, *6*, 32301.
74. Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and characterization of a peptide that enhances endosomal escape of delivered proteins in vitro and in vivo. *J Am Chem Soc* **2015**, *137* (44), 14084-93.
75. Falnes, P. O.; Sandvig, K., Penetration of protein toxins into cells. *Curr Opin Cell Biol* **2000**, *12* (4), 407-13.
76. Bade, S.; Rummel, A.; Reisinger, C.; Karnath, T.; Ahnert-Hilger, G.; Bigalke, H.; Binz, T., Botulinum neurotoxin type D enables cytosolic delivery of enzymatically active cargo proteins to neurones via unfolded translocation intermediates. *J Neurochem* **2004**, *91* (6), 1461-72.
77. Ballard, J. D.; Collier, R. J.; Starnbach, M. N., Anthrax toxin-mediated delivery of a cytotoxic T-cell epitope in vivo. *Proc Natl Acad Sci U S A* **1996**, *93* (22), 12531-4.

78. Liao, X.; Rabideau, A. E.; Pentelute, B. L., Delivery of antibody mimics into mammalian cells via anthrax toxin protective antigen. *Chembiochem* **2014**, *15* (16), 2458-66.
79. Rabideau, A. E.; Pentelute, B. L., Delivery of Non-Native Cargo into Mammalian Cells Using Anthrax Lethal Toxin. *ACS Chem Biol* **2016**, *11* (6), 1490-501.
80. Mechaly, A.; McCluskey, A. J.; Collier, R. J., Changing the receptor specificity of anthrax toxin. *MBio* **2012**, *3* (3).
81. Lofblom, J.; Feldwisch, J.; Tolmachev, V.; Carlsson, J.; Stahl, S.; Frejd, F. Y., Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. *FEBS Lett* **2010**, *584* (12), 2670-80.
82. McCluskey, A. J.; Olive, A. J.; Starnbach, M. N.; Collier, R. J., Targeting HER2-positive cancer cells with receptor-redirection anthrax protective antigen. *Mol Oncol* **2013**, *7* (3), 440-51.
83. Clarke, M. S.; McNeil, P. L., Syringe loading introduces macromolecules into living mammalian cell cytosol. *J Cell Sci* **1992**, *102* (Pt 3), 533-41.
84. Neumann, E.; Schaefer-Ridder, M.; Wang, Y.; Hofschneider, P. H., Gene transfer into mouse lymphoma cells by electroporation in high electric fields. *EMBO J* **1982**, *1* (7), 841-5.
85. McNeil, P. L.; Warder, E., Glass-Beads Load Macromolecules into Living Cells. *J Cell Sci* **1987**, *88*, 669-678.
86. Tsukakoshi, M.; Kurata, S.; Nomiya, Y.; Ikawa, Y.; Kasuya, T., A Novel Method of DNA Transfection by Laser Microbeam Cell Surgery. *Appl Phys B-Photo* **1984**, *35* (3), 135-140.
87. Hapala, I., Breaking the barrier: Methods for reversible permeabilization of cellular membranes. *Critical Reviews in Biotechnology* **1997**, *17* (2), 105-122.

88. Teng, K. W.; Ishitsuka, Y.; Ren, P.; Youn, Y. A.; Deng, X.; Ge, P. H.; Belmont, A. S.; Selvin, P. R., Labeling proteins inside living cells using external fluorophores for microscopy. *Elife* **2016**, *5*.
89. Sharei, A.; Zoldan, J.; Adamo, A.; Sim, W. Y.; Cho, N.; Jackson, E.; Mao, S.; Schneider, S.; Han, M. J.; Lytton-Jean, A.; Basto, P. A.; Jhunjhunwala, S.; Lee, J.; Heller, D. A.; Kang, J. W.; Hartoularos, G. C.; Kim, K. S.; Anderson, D. G.; Langer, R.; Jensen, K. F., A vector-free microfluidic platform for intracellular delivery. *Proc Natl Acad Sci U S A* **2013**, *110* (6), 2082-7.
90. Stewart, M. P.; Sharei, A.; Ding, X.; Sahay, G.; Langer, R.; Jensen, K. F., In vitro and ex vivo strategies for intracellular delivery. *Nature* **2016**, *538* (7624), 183-192.
91. Morisaki, T.; Lyon, K.; DeLuca, K. F.; DeLuca, J. G.; English, B. P.; Zhang, Z.; Lavis, L. D.; Grimm, J. B.; Viswanathan, S.; Looger, L. L.; Lionnet, T.; Stasevich, T. J., Real-time quantification of single RNA translation dynamics in living cells. *Science* **2016**, *352* (6292), 1425-9.
92. Rome, L. H.; Kickhoefer, V. A., Development of the Vault Particle as a Platform Technology. *Acs Nano* **2013**, *7* (2), 889-902.
93. Benner, N. L.; Zang, X.; Buehler, D. C.; Kickhoefer, V. A.; Rome, M. E.; Rome, L. H.; Wender, P. A., Vault Nanoparticles: Chemical Modifications for Imaging and Enhanced Delivery. *ACS Nano* **2017**, *11* (1), 872-881.
94. Seebeck, F. P.; Woycechowsky, K. J.; Zhuang, W.; Rabe, J. P.; Hilvert, D., A simple tagging system for protein encapsulation. *J Am Chem Soc* **2006**, *128* (14), 4516-7.
95. Worsdorfer, B.; Woycechowsky, K. J.; Hilvert, D., Directed evolution of a protein container. *Science* **2011**, *331* (6017), 589-92.

96. Sasaki, E.; Hilvert, D., Self-Assembly of Proteinaceous Multishell Structures Mediated by a Supercharged Protein. *J Phys Chem B* **2016**, *120* (26), 6089-95.
97. Sasaki, E.; Bohringer, D.; van de Waterbeemd, M.; Leibundgut, M.; Zschoche, R.; Heck, A. J.; Ban, N.; Hilvert, D., Structure and assembly of scalable porous protein cages. *Nat Commun* **2017**, *8*, 14663.
98. Tang, R.; Kim, C. S.; Solfiell, D. J.; Rana, S.; Mout, R.; Velazquez-Delgado, E. M.; Chompoosor, A.; Jeong, Y.; Yan, B.; Zhu, Z. J.; Kim, C.; Hardy, J. A.; Rotello, V. M., Direct Delivery of Functional Proteins and Enzymes to the Cytosol Using Nanoparticle-Stabilized Nanocapsules. *Acs Nano* **2013**, *7* (8), 6667-6673.
99. Mout, R.; Ray, M.; Tonga, G. Y.; Lee, Y. W.; Tay, T.; Sasaki, K.; Rotello, V. M., Direct Cytosolic Delivery of CRISPR/Cas9-Ribonucleoprotein for Efficient Gene Editing. *Acs Nano* **2017**, *11* (3), 2452-2458.
100. Wang, Y. A.; Yu, X.; Overman, S.; Tsuboi, M.; Thomas, G. J.; Egelman, E. H., The structure of a filamentous bacteriophage. *Journal of Molecular Biology* **2006**, *361* (2), 209-215.
101. DePorter, S. M.; McNaughton, B. R., Engineered M13 bacteriophage nanocarriers for intracellular delivery of exogenous proteins to human prostate cancer cells. *Bioconjug Chem* **2014**, *25* (9), 1620-5.

CHAPTER TWO

Mutagenesis Modulates the Uptake Efficiency, Cell-Selectivity, and Functional Enzyme Delivery of a Protein Transduction Domain²

2.1 Introduction

Previous work in the McNaughton laboratory identified a prostate cancer cell-selective protein transduction domain or cell penetrating peptide (CPP), which penetrates PC-3 prostate cancer cells with uptake efficiencies controlled by multivalency effects (**Figure 2.1**).¹ In order to assess the specific contribution each residue plays in cell uptake efficiency and cell-selectivity we completed alanine scanning mutagenesis of our reported prostate cancer cell-selective CPP, referred to as Ypep (N-YTFGLKTSFNVQ-C). Upon completion of these studies, we were able to identify two key residues. Extensive mutagenesis at these key residues generated multiple mutants with significantly improved uptake efficiency and ~4-fold improved cell selectivity for targeted cells. Our best mutant exhibited ~19-fold better uptake efficiency and ~4-fold improved cell-selectivity for a human prostate cancer cell line. In addition, while we could observe the delivery of functional fluorescent proteins to the interior of prostate cancer cells using Ypep, only modest functional enzyme delivery was achieved. In contrast, the most potent mutant, Ypep (G4N), was able to

² Adapted from: DePorter, S.M.; Lui, I.; Bruce, V.J.; Gray, M.A.; Lopez-Islas, M.; and McNaughton, B.R., *Mol. BioSystems*, **2013**, 10, 18. This work was led by a senior graduate student Sandra DePorter, and is a continuation of efforts for the identification of a protein transduction domain specific for prostate cancer cells (referred to as Ypep), and its further optimization. I assisted in gene construction, molecular cloning, and the identification an enhanced variant of Ypep that yields improved uptake and selectivity profiles, and achieves efficient delivery of an active enzyme.

deliver large quantities of a functional enzyme to the interior of human prostate cancer cells. Taken together, the research described below has significantly improved the efficiency, cell-selectivity, and functional utility of a prostate cancer CPP.

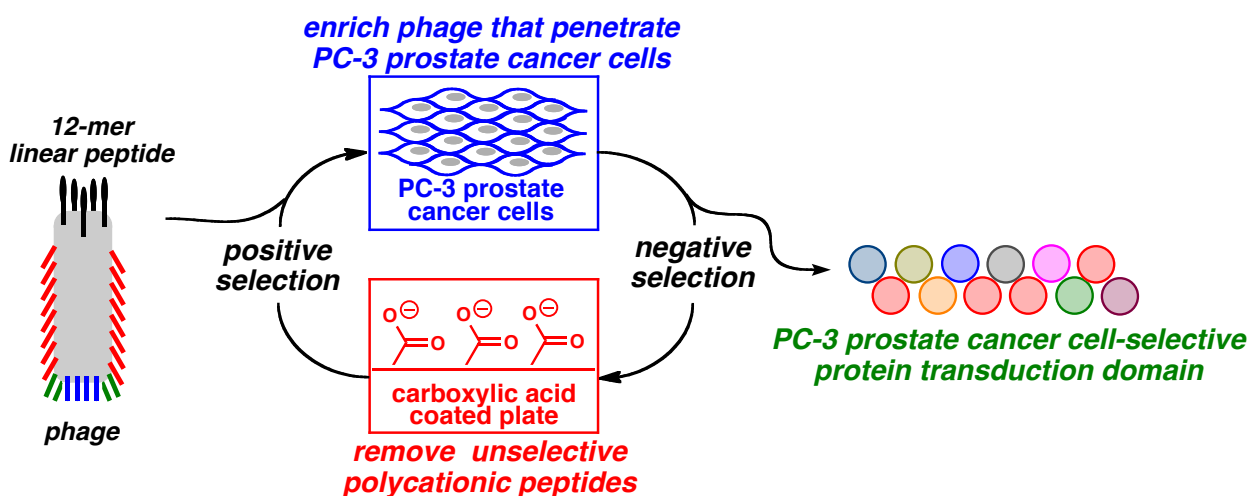


Figure 2. 1 Identification of Ypep. Evolution of a PC-3 prostate cancer cell-selective penetrating peptide. Phage are incubated with PC-3 prostate cancer cells enriched for cell penetration. Cell-penetrating phage are then incubated with a polyanionic tissue culture plate, and phage that don't bind the tissue culture plate are enriched. After three coupled rounds of positive and negative selections, the most abundant sequence enriched from this selection was "Ypep" (N-YTFGLKTSFNVQ-C). This figure was adapted from *Chemistry and Biology*, 2013, 20, 434.

As discussed in Chapter One, proteins are of great value as research tools, therapeutics, and imaging reagents.² The size and complexity of proteins can often endow these reagents with the ability to potently and selectively recognize disease-relevant macromolecules that confound drug discovery efforts limited to small molecules (<1,000 Da). In addition, a number of protein enzymes such as luciferase and horseradish peroxidase are commonly used for cellular imaging and basic research applications.³⁻⁴ In many ways, these reagents are ideally suited for imaging applications. For example, by virtue of their enzymatic activity, appreciable signal can be generated from a relatively small number of functional proteins. Despite clear advantages of protein-based approaches to therapy and cell imaging, the inability of most proteins to penetrate the lipid bilayer

membrane of mammalian cells largely limits the access of these reagents to cellular components displayed on the cell surface. The potential utility of functional proteins and enzymes with the ability to access the interior of mammalian cells has compelled researchers to address barriers to the intracellular delivery of these agents in mammalian cells. A more complete review of methods and technologies for intracellular protein delivery were discussed in Chapter One, but include electroporation,⁴ microinjection,⁵ liposomes,⁶ lipid-linked compounds,⁷ nanoparticles,⁸ fusions to receptor ligands,⁹ arginine grafting,¹⁰ supercharged proteins,¹¹⁻¹⁴ and cell penetrating peptides (CPPs)¹⁵⁻¹⁷.

In many ways, CPPs are ideally suited as protein delivery reagents. The proteinogenic amino acid composition and relatively small size (typically 7-20 L-amino acids) of these reagents allow researchers to express and purify soluble protein-CPP fusions with relative ease. A number of CPPs such as HIV-1 transactivator of transcription (Tat) peptide,¹⁸ the *Drosophila* Antennapedia-derived penetratin peptide,¹⁹ and polyarginine²⁰⁻²¹ have been used for intracellular protein delivery. However, the relatively modest uptake efficiency exhibited by some common CPPs, and a general lack of selectivity for diseased cells over healthy cells limits the full potential of these reagents. The development of new CPPs capable of both potent and cell-selective delivery of functional proteins, such as enzymes, to diseased cells would potentially expand the utility of protein-based approaches to cellular perturbation and imaging.

PC-3 cells are one of the most commonly used human prostate cancer cell lines in basic and translation research, and are useful in investigating the biochemical changes in advanced prostate cancer cells and in assessing their response to chemotherapeutic agents.²² Reagents that selectively recognize PC-3 cells are of potential value, since these cells do not express appreciable levels of prostate-specific membrane antigen (PSMA) – the most commonly used marker for prostate cancer

cell detection and targeted delivery. Therefore, PC-3 cells, or prostate cancer cells with a similar phenotype would likely evade targeted imaging or drug delivery strategies centered on PMSA recognition.

The work below describes our efforts to better understand the requirements for uptake of Ypep (**Figure 2.2A**) by PC-3 cells, in addition to its optimization in terms of potency and cell-selectivity. Starting with a significant number of Ypep mutants, these variants were assayed for their ability to deliver functional fluorescent protein or enzyme to PC-3 cells. We show that the best mutant delivers ~19-fold more fused protein to PC-3 cells than off-target non-cancer human embryonic kidney cells (HEK-293), with ~4-fold higher cell-selectivity, compared to native Ypep. In addition, the most potent and cell-selective mutant we identified delivers large quantities of functional enzyme to human prostate cancer cells. Similar to Ypep, uptake of the most potent and cell-selective mutant proceeds *via* energy-dependent endocytosis, suggesting that improvement in potency and cell-selectivity is not due to internalization *via* a different mechanism.

2.2 Alanine Scanning of Ypep Illuminates the Importance of Each Residue for PC-3 Cell-Specific Uptake

To assess the specific contribution each residue in Ypep plays in cell uptake efficiency, we made a library Ypep alanine mutants and expressed these peptides as N-terminal fusions to GFP. PC-3 cells were treated with 5 μ M of each Ypep-GFP mutant, a concentration previously shown to be sufficient for appreciable Ypep-GFP uptake. Cells were then exhaustively washed using conditions that we, and others, have previously shown to remove cell surface-bound protein. The amount of internalized GFP was measured by flow cytometry. As seen in **Figure 2.2B**, most mutations resulted in significantly lower GFP delivery, However, Ypep-Gly4Ala and Ypep-

Thr7Ala delivered ~3.8- and ~6.8-fold more GFP to PC-3 cells than native Ypep, respectively (**Figure 2.2B**). A number of commonly used CPPs such as Tat, polyarginine and penetration are polycationic, and rely on high-theoretical net charge for cell uptake. In contrast, Ypep has a theoretical net charge of +1. Interestingly, mutating the single positively charged (Lys6) to alanine decreased GFP uptake ~4-fold (**Figure 2.2B**). Based on these initial findings, we prepared a focused library of mutants with molecularly diverse residues at position four or seven.

2.3 Optimization of Ypep Mutants

Ypep mutants containing either negatively charged (aspartic acid), positively charged (lysine), aromatic (phenylalanine), hydrogen bond donating (serine), or amide (asparagine) functional group positions four or seven were expressed as N-terminal fusions to GFP. As seen in **Figure 2.2C**, the Gly4Asp mutant exhibited significantly lower uptake, and Gly4Phe and Gly4Ser mutants achieved only slightly higher uptake than Ypep-GFP. However, Gly4Lys and Gly4Asn mutants were significantly improved. Gly4Lys or Gly4 Asn mutants delivered ~3.2- and 19.2-fold more GFP to PC-3 cells, compared to native Ypep-GFP. Interestingly, small structural changes at position four significantly lowered uptake. While the surface receptor of Ypep and Ypep mutants is currently unknown, the fact that the addition of a methylene unit significantly lowers uptake supports a model wherein a well-defined interaction between Ypep-Gly4Asn and a cell surface receptor is required for efficient uptake, and not a less defined interaction that simply required sequence-defined functional group display on the CPP.

We next performed identical experiments to optimize residue seven. The Thr7Asp mutant exhibited essentially identical uptake efficiency as native Ypep. However, Thr7Lys, Thr7Ser, and Thr7Asn mutants all showed significantly lower transduction efficiencies. In contrast, Thr7Phe

mutant was significantly improved, as was observed to deliver ~7.6-fold more GFP to PC-3 cells, compared to native Ypep (**Figure 2.2D**). Based on this finding, we measured uptake efficiencies for Ypep variants containing all possible proteinogenic aromatic residues at position seven (**Figure 2.2D**). While both Thr7Tyr and Thr7Trp mutants significantly outperformed native Ypep, delivering ~6.8-fold and ~7.1-fold more GFP, respectively, neither outperformed the Thr7Phe mutant. In contrast, the Thr7His mutant showed significantly lower cell uptake compared to Thr7Tyr and Thr7Trp mutants, as well as native Ypep. Taken together, the reduced transduction we observed for the Thr7His and Thr7Lys mutants suggest that residues with positive charge, or partial positive charge, may not be tolerated at this position.

Combined synergistic effects can play important roles in many biological processes and macromolecular – substrate interactions. To assess if combinations of beneficial mutants are synergistic, we prepared three Ypep double mutants that contain combinations of the most beneficial single mutations at residues four and seven. Ypep double mutants that contain alanine, lysine, or asparagine at position four and phenylalanine at position seven were expressed N-terminal fusions to GFP and added to PC-3 cells as previously described. The Gly4Ala:Thr7Phe, Gly4Lys:Thr7Phe, and Gly4Asn:Thr7Phe double mutants were found to be ~3.5-, ~5.6-, and ~6.5-fold more efficient at GFP transduction than Ypep-GFP, respectively (**Figure 2.2E**). Interestingly, however, none of these double mutants exhibited higher uptake compared to the single mutant Ypep variants from which they were derived. Representative flow cytometry data are shown in **Figure S2.1**.

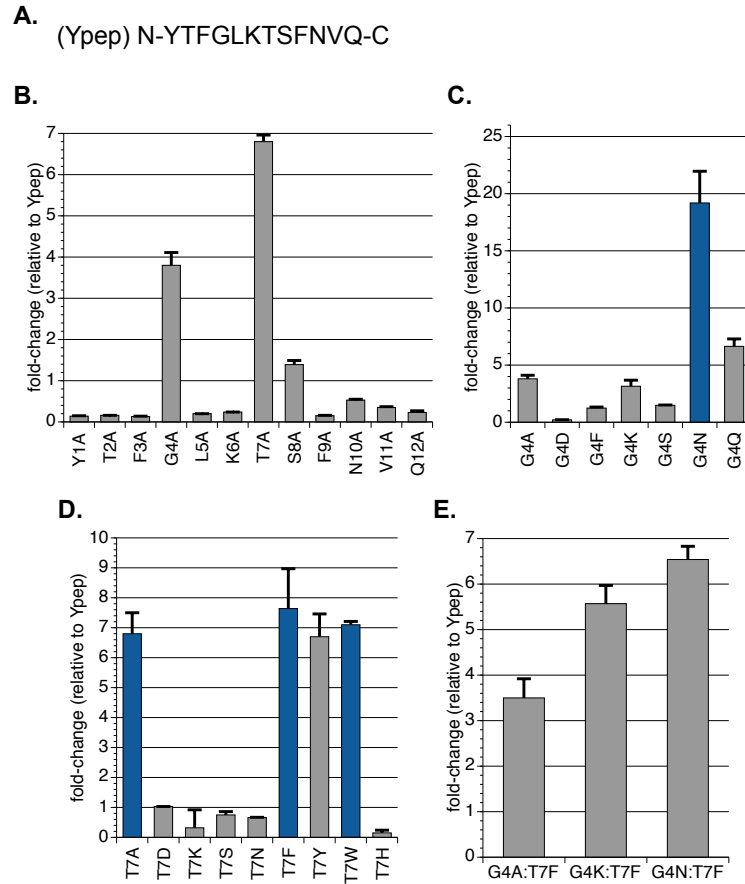


Figure 2. 2 Ypep ‘hot spot’ identification and optimization. (A) Amino acid sequence of Ypep. (B) Fold-change in GFP uptake for alanine mutants of Ypep-GFP, relative to Ypep-GFP. (C) Fold-change in GFP uptake for Ypep-GFP mutants at residue 4, relative to Ypep-GFP. (D) Fold-change in GFP uptake for Ypep-GFP mutants at residue 7, relative to Ypep-GFP. (E) Efficiency of GFP uptake for Ypep-GFP double mutants at residues 4 and 7. (B-E) In each example, PC-3 cells were treated with 5 μ M mutant Ypep-GFP, and then washed to remove cell surface-bound protein. GFP internalization was measured by flow cytometry. Values and error bars represent the mean and standard deviation of three independent experiments. Blue bars represent the four most active mutants.

2.4 Cytotoxicity and Mechanism of Uptake for Ypep Mutants

Based on these data, the Gly4Asn, Thr7Phe, Thr7, Trp, and Thr7Ala mutants (bars are colored blue in **Figure 2.2C** and **D**) were most improved over Ypep, with increased transduction efficiencies of ~19, ~8, and ~7-fold, respectively. To assess the cytotoxicity of Ypep variants under conditions required for appreciable uptake, we performed 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay (MTT) on PC-3 cells after treatment with 5 μ M Ypep(mutant)-GFP. These assays revealed no apparent cytotoxicity to PC-3 cells for any of the Ypep mutants

(**Figure S2.2**). GFP uptake was confirmed for these four mutants by live-cell fluorescence microscopy. While only a very small amount of internalized GFP was observed in PC-3 cells following treatment with 5 μ M Ypep-GFP (**Figure S2.3**), large amounts of internalized GFP was observed in cells following treatment with the same concentration of the four most active mutant Ypep-GFP fusions (**Figure 2.3**). Consistent with our flow cytometry data, the Gly4Asn mutant delivered the highest amount of GFP to the cell interior.

In the initial discovery of Ypep, we performed extensive mutagenesis to elucidate the mechanism of Ypep uptake. We found that while Ypep is taken up at 37 °C, it is not appreciably internalized with PC-3 cells incubated with a Ypep solution at 4 °C. This finding suggests that Ypep internalization proceeds *via* an energy-dependent endocytotic pathway. In order to determine if Ypep (Gly4Asn) uptake is consistent with the parent peptide, or if internalization proceeds *via* an alternative pathway, we incubated PC-3 cells with 1 μ M Ypep(Gly4Asn)-GFP at either 37 °C

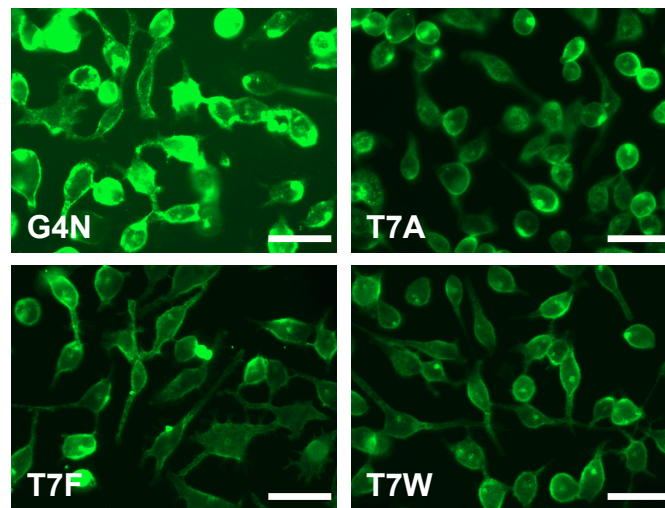


Figure 2. 3 Live cell fluorescence microscopy images of PC-3. Cells were treated with 5 μ M of the most efficient mutant Ypep-GFP fusions, then washed to remove cell surface-bound protein. Green color represents internalized GFP. Scale bar is 50 μ M. Lamp intensity was set at 50% with a 250 ms exposure for all images.

or 4 °C, washed cells to remove surface-bound material, and measured GFP internalization by microscopy and flow cytometry. Interestingly, similar to our findings for Ypep-GFP uptake, high levels of cell fluorescence is observed following treatment at 37 °C; however, no appreciable fluorescence is observed when cells are treated at 4 °C (**Figure S2.4**). Taken together, these findings suggest that like the parent peptide, the Ypep(Gly4Asn) mutant also relies on energy-dependent endocytosis for internalization.

2.5 Gly4Asn Ypep Outperforms Tat and Penetration Protein Transduction

To compare the uptake efficiencies of these reagents to commonly used CPPs, we treated PC-3 cells with solutions containing either Tat-GFP, penetratin-GFP, or the four best Ypep variants identified as a result of the previously discussed mutagenesis studies. As shown in **Figure 2.4**, following treatment with 1 μ M CPP-GFP fusion, and washing, all Ypep variants identified as a result of mutagenesis studies delivered significantly more GFP to the interior of PC-3 cells, compared to Tat-GFP fusions. Most notably, uptake efficiency in PC-3 cells treated with 1 μ M

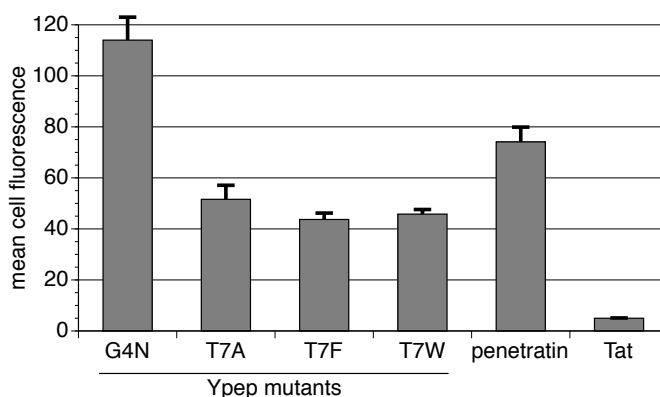


Figure 2. 4 Flow cytometry data comparisons of Ypep with penetratin and Tat. Data showing amounts of GFP delivered to PC-3 cells following treatment with 1 μ M Ypep(mutant)-GFP fusions, Tat-GFP fusion, or penetratin-GFP fusion then washed to remove cell surface bound protein. Values and error bars represent the mean and standard deviation of three independent experiments.

Ypep(Gly4Asn)-GFP was ~1.5-fold and 23-fold higher than cells treated with either penetratin-GFP or Tat-GFP fusions, respectively. Representative flow cytometry data is shown in **Figure S2.5**. Interestingly, we were unable to express appreciable amounts of soluble (Arg)₉-GFP, suggesting potential limitations to polyarginine-based approaches to protein delivery – at least in some cases.

2.6 Mutations Beneficial to Uptake Efficiency Also Increase the Cell-Selectivity of Protein Delivery

While the above mutational studies on Ypep resulted in numerous variants with improved transduction efficiency, the effect of those beneficial mutations on cell-selectivity was unclear. In order to assess the impact of these mutations on cell-selectivity, we compared CPP-GFP fusion uptake in PC-3 (target) and off-target non-cancer human embryonic kidney cells (HEK-293). Cells were treated with 0.1 – 1 μ M CPP-GFP fusion, washed as previously described to remove cell surface-bound material, and internalized GFP was measured by flow cytometry. As shown in **Figure 2.5**, a majority of the most efficient Ypep mutants also exhibited increased selectivity for PC-3 human prostate cancer cells. Representative flow cytometry data for **Figure 2.5** is shown in **Figure S2.6**. Consistent with our previous findings, Ypep delivered ~1.6-, ~1.8-, ~1.7-, or ~2.8-fold more GFP to PC-3 cells compared to HEK-293 cells, following treatment with 0.1, 0.25, 0.5, or 1.0 μ M solutions, respectively. While the Thr7Phe mutant exhibited similar selectivity for PC-3 cells (~1.6-, ~2.0-, ~2.6-, ~2.8-fold following 0.1 – 1.0 μ M treatment), the Gly4Asn, Thr7Trp, and Thr7Ala mutants were significantly more selective for PC-3 cells. For example, Gly4Asn, Thr7Trp, and Thr7Ala Ypep mutants were ~5.3, ~5.8-, and ~5.0-fold more selective for PC-3 prostate cancer cells compared to HEK-293 cells. Taken together, these studies demonstrate a

significant improvement in both the transduction efficiency and PC-3 cell-selectivity of multiple Ypep mutants found as a result of these studies.

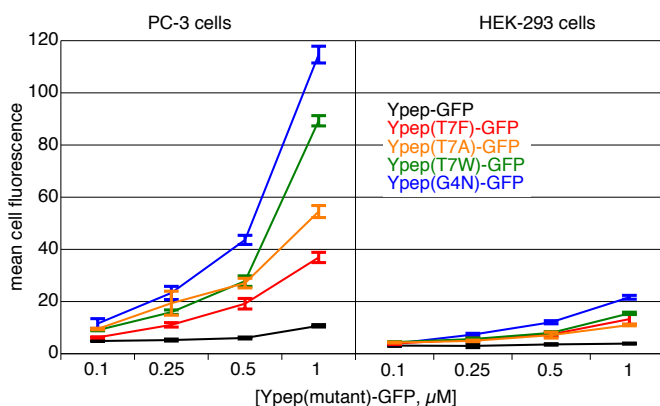


Figure 2. 5 Flow cytometry data showing the amount of internalized GFP in PC-3 cells or Hek-293 cells. Cells were treatment with 0.1, 0.25, 0.5, or 1.0 μM mutant Ypep-GFP, Ypep(T7F)-GFP, Ypep(T7A)-GFP, Ypep(T7W)-GFP, or Ypep(G4N)-GFP, then washed to remove cell surface-bound protein. Values and error bars represent the mean and standard deviation of three independent experiments.

2.7 Gly4Asn Ypep Mutant Delivers Functional Enzyme to PC-3 Cells

Perhaps the ultimate test of a CPP is the intracellular delivery of a functional enzyme. Luciferase is a class of enzymes that oxidize a photon-emitting substrate, resulting in bioluminescence. These enzymes enjoy extensive use as reporters and cell imaging reagents because of their high sensitivity, broad dynamic range, and operational simplicity.²³ NanoLuc luciferase (nLuc) is a recently reported variant of the small luciferase subunit from the deep-sea shrimp *Oplophorus gracilirostris*.²⁴ As a simple test for functional intracellular enzyme delivery, we measured luciferase activity in PC-3 cells following treatment with 1 μM nLuc, Ypep-nLuc, or Ypep(Gly4Asn)-nLuc. Consistent with the overwhelming majority of proteins, appreciable amounts of nLuc do not penetrate mammalian cells. Cells treated with nLuc, washed to remove surface-bound protein, and treated with furimazine (nLuc specific substrate), exhibited very little luminescence (**Figure 2.6**). Similar to our previous findings, relatively modest functional enzyme

delivery was achieved *via* Ypep-dependent delivery. Cells treated with Ypep-nLuc, then washed as described above, were ~6.1-fold more luminescent than cells treated with nLuc alone (**Figure 2.6**). In contrast, cells similarly treated with Ypep(Gly4Asn)-nLuc were ~42-fold more luminescent than cells treated with nLuc. These findings suggest that relatively large amounts of enzymatically active Ypep (Gly4Asn)-nLuc were delivered to the interior of PC-3 cells (**Figure 2.6**). Importantly, cells are not lysed at any point during the luciferase assay. Therefore, luminescence generated during these experiments is the action of active nLuc enzyme within the cell interior. **Figure S2.7** shows PAGE analysis of all proteins used in this work.

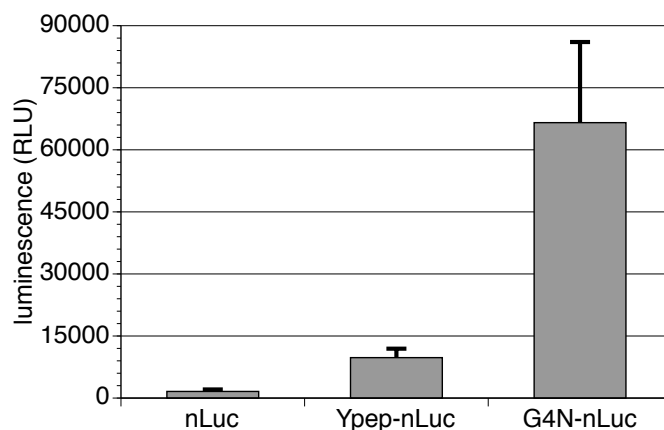


Figure 2. 6 Efficiency of nanoluciferase (nLuc) delivery to human prostate cancer cells (PC-3). PC-3 cells were treated with either nLuc or Ypep-nLuc or Ypep(Gly4Asn)-nLuc, then washed to remove cell surface-bound protein. Functional nLuc and Ypep-nLuc does not appreciably penetrate PC-3 cells; however, relatively high levels of internalized functional nLuc are observed in cells following treatment with Ypep(G4N)-nLuc.

2.8 Conclusion

In summary, we recently reported Ypep, a prostate cancer cell-selective CPP, with uptake efficiency and cell-selectivity profiles that are dependent on multivalency effects. When a single

copy of Ypep is fused to GFP, modest uptake efficiency and cell-selectivity is observed. Fusion of a single Ypep peptide to the N-terminus of nLuc does not deliver appreciable functional protein to PC-3 cells. Mutational studies have revealed a number of Ypep variants with significantly improved protein transduction efficiency and selectivity for PC-3 human prostate cancer cells. Amazingly, a single mutation to Ypep (Gly4Asn) resulted in a variant with ~19.2-fold higher transduction efficiency and ~4-fold better selectivity for PC-3 cells over off-target HEK-293 cells (**Figure 2.6**). In contrast to Ypep, Ypep(Gly4Asn) delivered appreciable levels of nLuc to the interior of PC-3 cells. Taken together, the findings described in this paper significantly improve the functional utility of Ypep-dependent delivery of exogenous proteins to the interior of PC-3 prostate cancer cells. Our data suggest that the Ypep mutants described here are well suited to serve as reagents for PC-3 cell-selective delivery of imaging and enzymatic proteins for basic research and biomedical applications.

REFERENCES

1. DePorter, S. M.; Lui, I.; Mohan, U.; McNaughton, B. R., A protein transduction domain with cell uptake and selectivity profiles that are controlled by multivalency effects. *Chem Biol* **2013**, *20* (3), 434-44.
2. Carter, P. J., Introduction to current and future protein therapeutics: a protein engineering perspective. *Exp Cell Res* **2011**, *317* (9), 1261-9.
3. Li, X.; Nakajima, Y.; Niwa, K.; Viviani, V. R.; Ohmiya, Y., Enhanced red-emitting railroad worm luciferase for bioassays and bioimaging. *Protein Sci* **2010**, *19* (1), 26-33.
4. Yan, M.; Du, J.; Gu, Z.; Liang, M.; Hu, Y.; Zhang, W.; Priceman, S.; Wu, L.; Zhou, Z. H.; Liu, Z.; Segura, T.; Tang, Y.; Lu, Y., A novel intracellular protein delivery platform based on single-protein nanocapsules. *Nat Nanotechnol* **2010**, *5* (1), 48-53.
5. Barsagi, D.; Feramisco, J. R., Microinjection of the Ras Oncogene Protein into Pc12 Cells Induces Morphological-Differentiation. *Cell* **1985**, *42* (3), 841-848.
6. Gregoriadis, G., Engineering liposomes for drug delivery: progress and problems. *Trends Biotechnol* **1995**, *13* (12), 527-37.
7. Zelphati, O.; Wang, Y.; Kitada, S.; Reed, J. C.; Felgner, P. L.; Corbeil, J., Intracellular delivery of proteins with a new lipid-mediated delivery system. *J Biol Chem* **2001**, *276* (37), 35103-10.
8. Utama, R. H.; Guo, Y.; Zetterlund, P. B.; Stenzel, M. H., Synthesis of hollow polymeric nanoparticles for protein delivery via inverse miniemulsion periphery RAFT polymerization. *Chem Commun (Camb)* **2012**, *48* (90), 11103-5.

9. Gabel, C. A.; Foster, S. A., Mannose 6-phosphate receptor-mediated endocytosis of acid hydrolases: internalization of beta-glucuronidase is accompanied by a limited dephosphorylation. *J Cell Biol* **1986**, *103* (5), 1817-27.
10. Fuchs, S. M.; Raines, R. T., Arginine grafting to endow cell permeability. *ACS Chem Biol* **2007**, *2* (3), 167-70.
11. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R., A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. *Chem Biol* **2011**, *18* (7), 833-8.
12. Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* **2010**, *5* (8), 747-52.
13. McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R., Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc Natl Acad Sci U S A* **2009**, *106* (15), 6111-6.
14. Thompson, D. B.; Villasenor, R.; Dorr, B. M.; Zerial, M.; Liu, D. R., Cellular uptake mechanisms and endosomal trafficking of supercharged proteins. *Chem Biol* **2012**, *19* (7), 831-43.
15. Deshayes, S.; Morris, M. C.; Divita, G.; Heitz, F., Cell-penetrating peptides: tools for intracellular delivery of therapeutics. *Cell Mol Life Sci* **2005**, *62* (16), 1839-49.
16. Mae, M.; Langel, U., Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery. *Curr Opin Pharmacol* **2006**, *6* (5), 509-14.
17. Morris, M. C.; Deshayes, S.; Heitz, F.; Divita, G., Cell-penetrating peptides: from molecular mechanisms to therapeutics. *Biol Cell* **2008**, *100* (4), 201-17.

18. Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **1988**, *55* (6), 1179-88.
19. Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A., The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem* **1994**, *269* (14), 10444-50.
20. Dubikovskaya, E. A.; Thorne, S. H.; Pillow, T. H.; Contag, C. H.; Wender, P. A., Overcoming multidrug resistance of small-molecule therapeutics through conjugation with releasable octaarginine transporters. *Proc Natl Acad Sci U S A* **2008**, *105* (34), 12128-33.
21. Fuchs, S. M.; Raines, R. T., Polyarginine as a multifunctional fusion tag. *Protein Sci* **2005**, *14* (6), 1538-44.
22. Kaighn, M. E.; Narayan, K. S.; Ohnuki, Y.; Lechner, J. F.; Jones, L. W., Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest Urol* **1979**, *17* (1), 16-23.
23. Thorne, N.; Inglese, J.; Auld, D. S., Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology. *Chem Biol* **2010**, *17* (6), 646-57.
24. Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V., Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7* (11), 1848-1857.

CHAPTER THREE

Resurfaced Cell-Penetrating Nanobodies: A Potentially General Scaffold for Intracellularly Targeted Protein Discovery³

3.1 Introduction

As discussed prior, proteins can often recognize and modulate disease-relevant macromolecules that present a challenge to small-molecule reagents by virtue of their size, functional group diversity and complex structure. Additionally, high-throughput screening and evolution-based methods often make the discovery of new protein binders simpler than the analogous small-molecule discovery process. However, as thoroughly discussed in Chapter One, most proteins do not cross the lipid bilayer membrane of mammalian cells. This largely limits the scope of protein therapeutics and basic research tools to those targeting disease-relevant receptors on the cell surface or extracellular matrix. A common theme for endowing cellular uptake is to increase the theoretical net charge of the protein through cationic resurfacing of the desired protein. However, in our experience, many proteins are not amenable to such extensive mutagenesis. Here, we report that nanobodies – a small and stable protein that can be evolved to recognize virtually any disease-relevant receptor – are amenable to cationic resurfacing, which results in cell internalization. Once internalized, these nanobodies access the cytosol. Polycationic resurfacing

³ Adapted from: Bruce, V.J.; Lopez-Islas, M.; and McNaughton, B.R., *Protein Science*, **2016**, 25, 1129. I led this effort for the delivery of polycationic resurfaced nanobodies to the cytosol of mammalian cells. Monica Lopez-Islas, an undergraduate research assistant, made contributions to the cloning and purification of proteins used in this work.

does not appreciably alter the structure, expression, and function (target recognition) of a previously reported GFP-binding nanobody, and multiple nanobody scaffolds are amenable to polycationic resurfacing. Given this, we propose that polycationic resurfaced cell-penetrating nanobodies might represent a general scaffold for intracellularly targeted protein drug discovery.

One cannot overstate the unique opportunities proteins offer as therapeutics and basic research tools. While all small-molecules reported to date modulate a very small percentage of the proteome (~2%) – and only a handful of protein structural classes – the size, functional group diversity, and complex three-dimensional structure of proteins can enable much broader recognition.¹ Moreover, various high-throughput screening and evolution-based methods make the discovery of new protein binders simpler than the analogous small-molecule focused process.²⁻⁴

A forefront challenge to the broader use of proteins in biomedical applications is their general inability to efficiently cross the lipid bilayer of mammalian cells and access the cytosol. Thus, most current protein drugs and basic research tools target disease-relevant receptors that reside on the surface of the cell or the extracellular matrix. Efforts to unlock the full potential of proteins in biomedical applications by enabling potent and functional cell penetration have been a major focus of modern biologics research.⁵⁻⁹ Incorporation of polycationic linkages – such as polyarginine – has previously been described as a means to enable cell penetration of various cargo, including proteins.¹⁰ More recently, researchers have used protein engineering to generate polycationic features on the protein surface. For example, Raines and coworkers reported that “arginine grafting” – mutagenesis of clustered solvent exposed amino acids to arginine – enables cellular uptake.¹¹ In a conceptually similar strategy, Liu and coworkers have shown that protein “supercharging” – extensive mutagenesis of a large number of solvent-exposed residues to

positively charged lysine or arginine as discussed in Chapter One – results in potent penetration of mammalian cells.¹²⁻¹⁴

While these polycationic resurfacing methods endow potent cell penetration, a major challenge to their broader application is the lack of established and broadly applicable guidelines for this extensive mutagenesis. Relatively little is known about how to dramatically resurface a protein with a polycationic feature in a manner that does not dramatically alter or abolish its utility and/or function (stability, target affinity, expression in *Escherichia coli*). In our experience, even structurally similar proteins respond differently to such extensive mutagenesis, and many proteins of therapeutic interest were not amenable to polycationic resurfacing. Perhaps a simpler approach is to focus effort on developing a single resurfaced polycationic, cell-penetrating, protein scaffold that is stable, expresses in *E.coli*, maintains the function of the original protein, but can be evolved to bind virtually any disease-relevant intracellular target.

Single-domain antibodies derived from camelids, referred to as nanobodies (**Figure 3.1**), have emerged as an alternative to traditional antibodies.¹⁵⁻¹⁶ Features of nanobodies make them well-suited as a general scaffold for protein drug discovery and polycationic resurfacing. In contrast to monoclonal antibodies, nanobodies are produced in large amounts in bacterial expression systems, are small in size (~15 kDa), are usually very stable and often bind their target with excellent affinities ($K_D \sim 1 - 100$ nM) through interactions involving well-defined antigen binding loops, referred to as the complementarity-determining regions (CDRs, **Figure 3.1A**, purple).¹⁷⁻¹⁸ Maturation of amino acids within one or more of the CDR loops by high-throughput screening or evolution-based methods can lead to new proteins that bind virtually any macromolecular target.¹⁹⁻

The structure of nanobodies is highly homologous, and consists of the CDRs, where the nanobody recognizes its target, and a framework region that is rich in β -sheet and loop structure. Researchers have shown that CDR loops can be inserted in a particularly stable nanobody scaffold, resulting in a new nanobody that retains affinity (based on choice of CDR loops), but has improved stability (based on judicious choice of the nanobody framework).²² Based on this observation, we hypothesized that if we could engineer polycationic resurfaced nanobody scaffolds, the resulting framework region could likely serve as a generic scaffold for the discovery of cell-penetrating nanobodies that bind and modulate disease-relevant intracellular receptors.

3.2 Polycationic Resurfacing of Three Previously Reported Nanobody Frameworks

We began with a previously reported nanobody that binds the Green Fluorescent Protein (GFP, **Figure 3.1A**).¹⁷⁻¹⁸ Structural analysis of the nanobody that binds GFP (referred to as NB1, herein) revealed a large solvent-exposed surface consisting of a β -sheet and loop structure – within the framework region – that is distinct from the CDR loops. We hypothesized that extensive polycationic resurfacing within this region by mutation of a critical number of residues to arginine (R) or lysine (K) (**Figure 3.1B**, blue spheres) should endow cell penetration. The sequence and mutagenesis of the wild-type nanobody and resurfaced variant is shown in **Figure 3.2A**. Satisfyingly, the polycationic resurfaced GFP-binding nanobody (referred to as pcNB1, herein), which has a theoretical net charge of +14, expresses as a soluble protein in *E. coli* (**Figure 3.2B**). Expanding on this successful result, we performed analogous polycationic resurfacing on two other recently reported nanobodies, which bind HER2²¹ or β -lactamase²³, respectively (referred to as NB2 or NB3, herein). The sequence of the wild-type nanobodies and resurfaced variants is shown in **Figure 3.2A**. While the size and sequence of the CDR loops differ extensively, and small

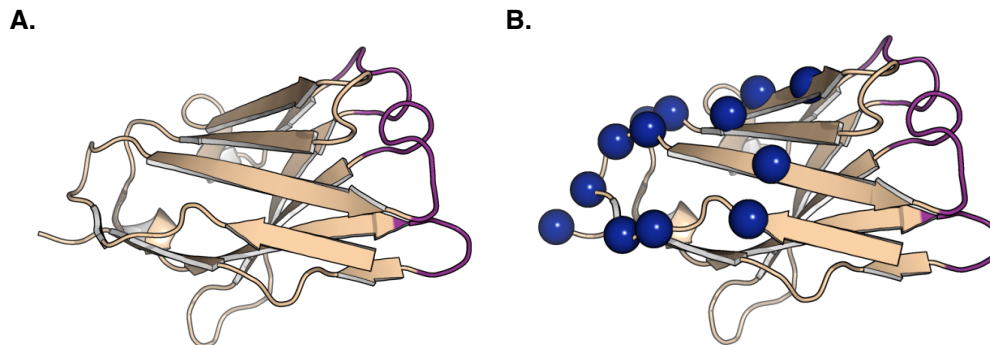


Figure 3. 1 Typical protein structure for nanobodies. (A) A previously reported nanobody that binds Green Fluorescent Protein (GFP), PDB: 3OGO. This nanobody is referred to as NB1 in this work. Complementarity-determining region (CDR) loops are highlighted in purple. (B) Residues on NB1 that were mutated to either arginine or lysine to generate the resurfaced polycationic nanobody (pcNB1) are highlighted with blue spheres.

changes in the framework sequence of the wild-type nanobody exist, the resulting polycationic resurfaced nanobodies (referred to as pcNB2 or pcNB3, herein), which have a theoretical net charge of +14 and +15, respectively, express in *E. coli* as soluble proteins (**Figure 3.2B** and **Figure S3.1**). Our resurfacing design is summarized as follows: First, we set a goal of generating nanobodies with a theoretical net charge of approximately +15, based on previous cell-penetrating studies on supercharged or arginine grafted GFP's.^{11, 13, 24} Second, we focused our mutation on residues that were well within the framework region, and not in or near the CDR loops. Third, we tried to space out mutations, so as to avoid cation/cation repulsion, which would likely effect protein folding and/or stability. Once candidate residues were identified, based on the above criteria, we considered whether a mutation should result in installation of an arginine or lysine. Since arginine results in better cell surface binding, and cell-penetration,¹⁰ compared to lysine, we favored mutation to arginine, unless the size of neighboring residues suggested that mutation to the relatively large arginine would potentially result in steric clashing. Interestingly, given this relatively simplistic resurfacing design, our initial attempt at polycationic resurfacing was

successful for all three nanobody scaffolds. Since analogous attempts to resurface other protein scaffolds are often unsuccessful, in our experience, we conclude that nanobodies may be particularly amenable to such polycationic resurfacing.

3.3 Polycationic Resurfacing Does Not Alter Structure, but does Endow Internalization of Mammalian Cells

We next assessed structural features of the wild-type and resurfaced nanobodies by circular dichroism. All nanobodies examined – wild-type and resurfaced variants – have a circular dichroism spectra similar to a previously reported nanobody²⁵ (**Figure 3.2C**). Collectively, expression of all resurfaced proteins in a soluble form, and similarities in the circular dichroism spectra of the wild-type and mutated variants, suggest that no dramatic structural changes occur as a result of polycationic resurfacing.

To determine uptake efficiency we fused each of our polycationic to GFP and measured uptake by flow cytometry. 3T3 cells were first treated with 10-500 nM polycationic resurfaced nanobody-GFP fusion, then washed with a phosphate buffered saline solution containing 20 U/mL heparin sulfate – which has been previously shown to remove cell surface bound protein especially supercharged protein.^{6-7, 12-14, 24} Following treatment with trypsin, which has also shown to remove and/or degrade surface bound protein,²⁶ intracellular levels of nanobody-GFP was measured by flow cytometry. For each resurfaced nanobody we observed a concentration-dependent increase of internalized fusion protein, as seen in **Figure 3.3A-C**. In contrast, fusion proteins composed of the wild-type protein and GFP do not appreciably penetrate 3T3 cells (**Figure 3.3A-C**). Internalization was further analyzed by fluorescence microscopy (**Figure 3.3D-F** and **Figure**

S3.2). Significant levels of each resurfaced nanobody-GFP fusion protein were observed in 3T3 cells, following the above described washing conditions to remove cell surface-bound protein.

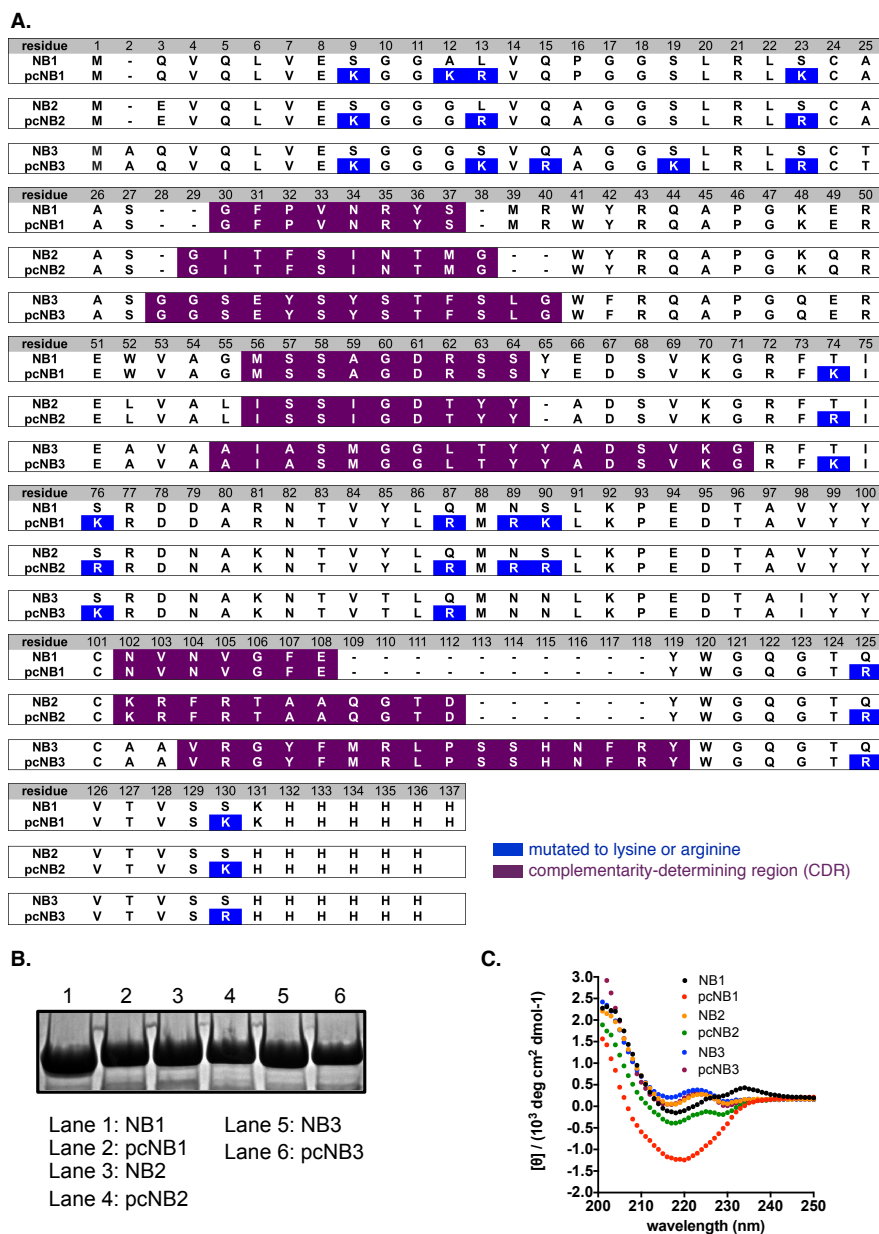


Figure 3. 2 Resurfacing and purification of polycationic nanobodies. (A) Sequence of wild-type nanobodies (NB1-3) and resurfaced polycationic nanobodies (pcNB1-3) described in this work. (B) PAGE analysis of wild-type and resurfaced polycationic nanobodies described in this work. (C) Circular dichroism spectra of wild-type (NB1-3) and resurfaced polycationic nanobodies (pcNB1-3) described in this work.

3.4 Polycationic Resurfaced Nanobodies Access the Cytosol of Mammalian Cells

Interestingly, internalized arginine grafted GFP and supercharged GFP appear as punctate foci in fluorescence microscopy images^{11, 14} (see **Figure S3.3**) – suggesting encapsulation within endosomes. However, the resurfaced nanobody-GFP fusions do not appear as such, suggesting that appreciable amounts of these internalized nanobodies might access the cytosol. This is critical, since the discovery of future cell-penetrating nanobodies based on these scaffolds would need to access the cytosol in order to engage therapeutically-relevant intracellular targets. This important aspect of cell uptake was further analyzed using a previously described method.²⁷⁻²⁸ 3T3 cells were first treated with 250 nM nanobody-GFP or polycationic resurfaced nanobody-GFP fusions, then washed as described above to remove cell surface bound protein. Cells were then lysed with a solution containing digitonin – which breaks the cell surface lipid bilayer, but not endosomes. The cellular location of each internalized fusion protein (cytosolic or endosomal) was then assessed by Western blot, using as anti-GFP antibody (a marker for internalized fusion protein), anti-Erk 1/2 antibody (a marker for the cytosol) or anti-Rab5 antibody (a marker for endosomes). No appreciable amount of GFP or wild-type nanobody-GFP fusion is found within the cytosolic extraction (following cell lysis with digitonin, **Figure 3.3G**, lanes 2-5), and no appreciable amount of Rab5 is observed (indicating that the lysis does not contain broken endosomes). In contrast, internalized resurfaced nanobody-GFP fusions appear in the cytosol – in the fraction that tests positive for the cytosolic marker Erk 1/2 but does not have any appreciable amount of the endosome marker Rab5 (**Figure 3.3G**, lanes 6-8). Thus, the polycationic resurfaced protein is capable of dragging another protein (GFP) into the cytosol of a mammalian cell.

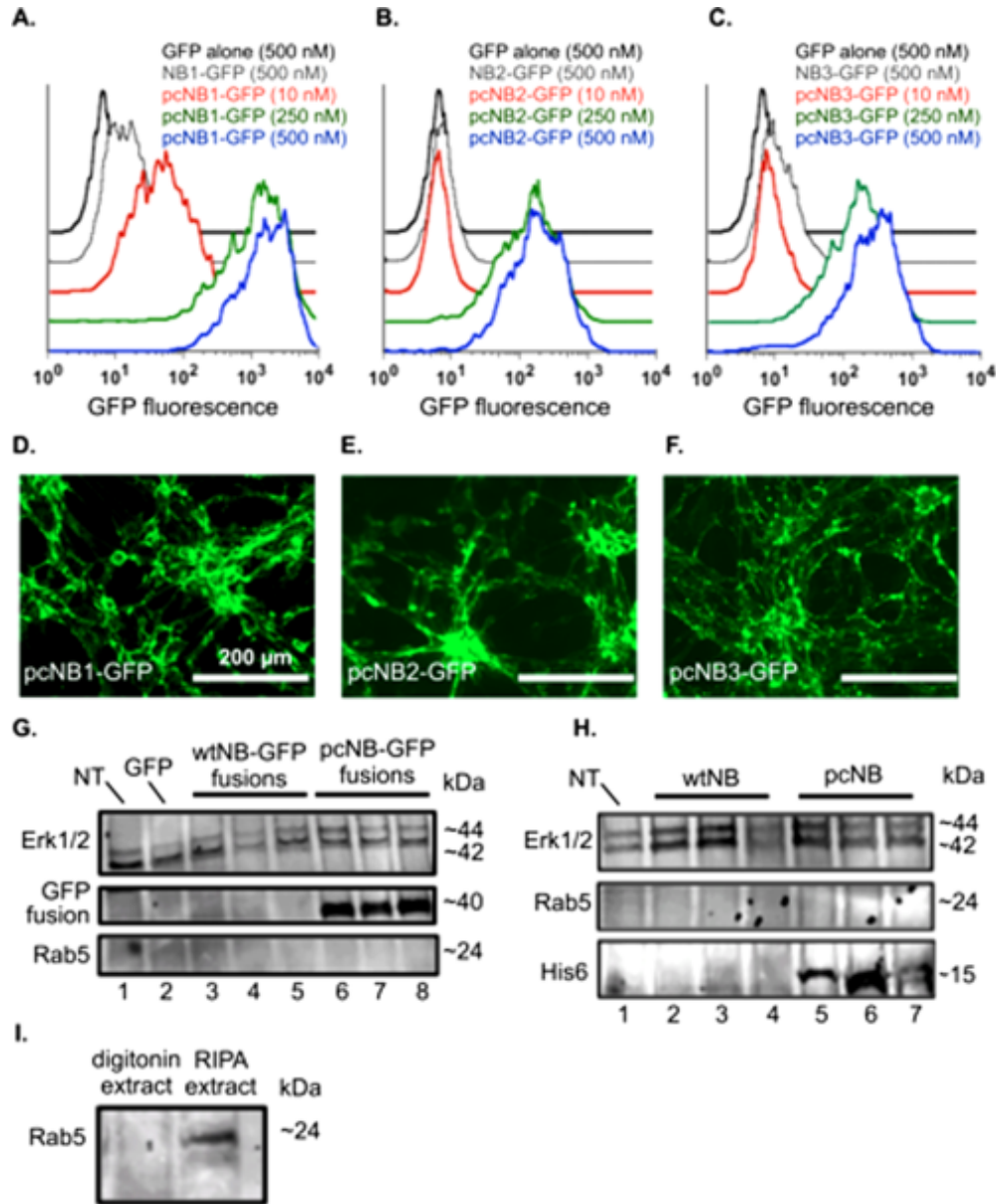


Figure 3. 3 Polycationic Resurfaced Nanobodies Reach the Cytosol of Mammalian Cells. (A-C) Flow cytometry data that supports concentration-dependent uptake of resurfaced polycationic nanobody-GFP fusion proteins, but not GFP alone (black line) or wild-type nanobody-GFP fusion (gray line). Red line = 10 nM treatment; green line = 250 nM treatment; blue line = 500 nM treatment. (D-F) Fluorescence microscopy images of 3T3 cells following treatment with 250 nM resurfaced nanobody-GFP fusions. (G) Western blot analysis of digitonin cell lysate for Erk1/2 (cytosolic marker), GFP (internalized resurfaced nanobody-GFP fusion protein), or Rab5 (endosome marker). Lane 1 = no treatment; lane 2 = wild-type GFP; lane 3 = wild-type NB1-GFP fusion; lane 4 = NB2-GFP fusion; lane 5 = NB3-GFP fusion; lanes 6-8 = polycationic resurfaced nanobody-GFP fusions analogous to lanes 3-5. (H) Western blot analysis of digitonin cell lysate for Erk1/2 (cytosolic marker), His₆ (internalized resurfaced nanobody), or Rab5 (endosome marker). Lane 1 = no treatment; lane 2 = wild-type NB1; lane 3 = NB2; lane 4 = NB3; lanes 5-7 = polycationic resurfaced NB1, NB2, or NB3, respectively. (I) Western blot showing no Rab5 (endosome marker) in cell lysate following digitonin lysis, but in extract following RIPA lysis. For all figures, experiments were run in triplicate and representative data are shown.

Fusion to GFP is not required for polycationic resurfaced nanobodies to penetrate mammalian cells and access the cytosol. The same assay was repeated, but cells were incubated with 500 nM wild-type nanobodies or polycationic resurfaced nanobodies with a minimal His₆ tag (for purification and identification by Western blot). As before, wild-type nanobodies do not appear in the cytosolic fraction (**Figure 3.3H**, lanes 2-4). In contrast, we observe appreciable levels of the His₆ labeled polycationic resurfaced nanobodies in the cytosolic portion of cell lysate (which stains for cytosolic markers Erk 1/2, **Figure 3.3H**, lanes 5-7). As a positive control, when cells are lysed with RIPA buffer, which breaks apart the lipid bilayers of both the cell surface and endosomes, we observe the protein marker for endosomes (Rab5, **Figure 3.3I**).

3.5 Polycationic Resurfacing Does Not Alter Nanobody Function and Stability

Having established that the resurfaced nanobodies penetrate mammalian cells and accesses the cytosol, we next explored if this extensive mutagenesis alters function (compared to the wild-type nanobody). This is important, since we want to endow cell penetration, but maintain a structure capable of binding a target (ultimately an intracellular target following CDR affinity maturation). Among the set of starting nanobodies, retention of function is most easily assessed using the GFP-binding nanobody¹⁷⁻¹⁸, since its binding partner (GFP) is easily expressed and observed, and this interaction is particularly well characterized. In order to determine if polycationic resurfaced GFP-binding nanobody (pcNB1 in this work) still binds GFP in a living cell, we co-expressed His₆-labeled NB1 or pcNB1 and untagged GFP in *E. coli* from a pET-DUET plasmid. Following purification on nickel-NTA resin, purified proteins were analyzed by polyacrylamide gel electrophoresis (PAGE) and Coomassie staining. Unsurprisingly, untagged GFP co-purifies with His₆-NB1 (**Figure 3.4A**, lane 2). Gratifyingly, untagged GFP also co-purified with the

polycationic resurfaced variant His₆-pcNB1, suggesting that GFP affinity is retained, even in the chemically complex environment of a living cell (*E. coli*). In addition, we performed Isothermal Titration Calorimetry (ITC) experiments to measure affinity between NB1 or pcNB1 and GFP. Because of the high affinity ($K_D \sim 1$ nM), it is difficult to use ITC to measure the equilibrium binding constant with precision. However, as previously reported¹⁷ NB1 binds GFP with essentially the same affinity (**Figure S3.4**).

We next set out to determine how polycationic resurfacing affects protein stability and robustness – important features when considering proteins as basic research tools and therapeutic leads. As previously stated, nanobodies are highly stable and robust proteins. Previous reports have shown that some nanobodies – including the GFP-binding nanobody – can be thermally denatured, but refolded when cooled slowly. To see if the polycationic resurfaced GFP-binding nanobody (pcNB1) has the same level of stability and robustness, we tested its ability to recover from thermal denaturation. Both the wild-type His₆-labeled GFP-binding nanobody (His₆-NB1) and His₆-labeled polycationic resurfaced variant (His₆-pcNB1) were heated to 100 °C for 2 minutes, then allowed to cool to room temperature over the course of 2 hours. After cooling, the samples were incubated with cell lysate from *E.coli* that expressed recombinant GFP lacking a His₆ label. This solution was then incubated with nickel-NTA resin, the resin was washed, and nickel-bound protein was eluted with imidazole solution. Under these conditions, if His₆-NB1 or His₆-pcNB1 recover from thermal denaturation and regain function (GFP affinity), elution from the column should include both NB1 or pcNB1 and bound GFP. Eluted solutions were analyzed by a long wave (365 nm) hand-held lamp for the presence of GFP. As expected, no appreciable GFP fluorescence is seen when illuminating elution from nickel-bound His₆-NB1 and His₆-pcNB1 (**Figure 3.4B**, tubes 1-2). However, GFP fluorescence (indicating co-elution of the His₆-nanobody

and bound GFP) is observed in elution from nickel-bound His₆-NB1 and untagged GFP (**Figure 3.4B**, tubes 3-4). As might be expected with supercharged variants, similar levels of GFP fluorescence is observed in elution from nickel-bound His₆-pcNB1 and untagged GFP (**Figure 3.4B**, tubes 5-6). As a positive control, elution from nickel-bound His₆-GFP is similarly fluorescent (**Figure 3.4B**, tube 7). Collectively, these data show the nanobodies NB1 and pcNB1 are not appreciably fluorescent, and NB1 and pcNB1 are not appreciably fluorescent, and NB1 and pcNB1 are able to recover from thermal denaturation and bind GFP. Thus, polycationic resurfacing does not appreciably alter protein nanobody stability and robustness. Unsurprisingly, GFP lacking a His₆ tag does not have appreciable affinity for nickel-NTA (**Figure 3.4B**, tube 8).

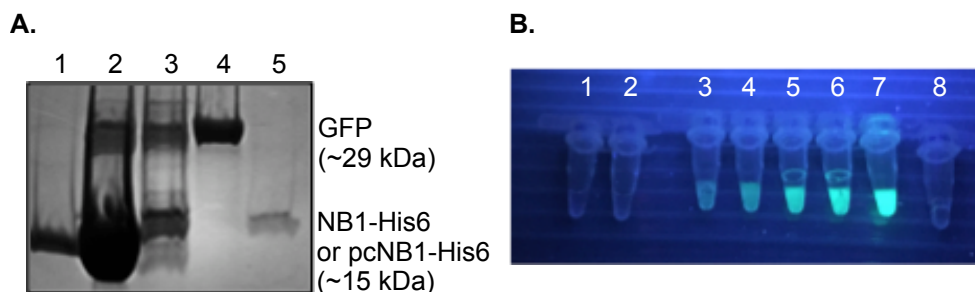


Figure 3. 4 Polycationic resurfaced nanobodies retain function and stability. (A) Lane 1: His₆-NB1; Lane 2: co-purification of untagged GFP with His₆-NB1 from *E. coli* cell lysate; Lane 3: co-purification of untagged GFP with His₆-pcNB1; Lane 4: His₆-GFP; Lane 5: His₆-pcNB1. (B) Tube 1: His₆-NB1; Tube 2: His₆-pcNB1; Tubes 3-4: His₆-NB1 and co-eluted GFP; Tubes 5-6: His₆-pcNB1 and co-eluted GFP; Tube 7: His₆-GFP; Tube 8: untagged GFP. For all figures, experiments were run in triplicate and representative data are shown.

3.6 Conclusion

In conclusion, the inability of most proteins to penetrate mammalian cells greatly limits the identification of new protein therapeutics that bind and modulate disease-relevant intracellular targets. Proteins with engineered solvent-exposed cationic features penetrate mammalian cells, but

a lack in general guidelines for such extensive mutagenesis on a number of therapeutically-relevant proteins, limits the broader application of this approach. An alternative strategy is to identify a single protein scaffold that is amenable to polycationic resurfacing, is cell-penetrating, access the cytosol of mammalian cells, and can be evolved using *in vitro* and *in vivo* techniques to generate cell-penetrating proteins that bind and modulate intracellular disease-relevant targets. Here, we show that three previously reported nanobodies can be resurfaced to display an extended polycationic feature on the framework region. This mutagenesis results in a new nanobody that is potently cell-penetrating, while maintaining structure, function and stability/robustness. Based on these findings, we anticipate that polycationic resurfaced nanobodies might serve as a general scaffold for the discovery of protein basic research tools and therapeutic leads that target disease-relevant intracellular receptors.

REFERENCES

1. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L., How many drug targets are there? *Nat Rev Drug Discov* **2006**, 5 (12), 993-6.
2. Smith, G. P.; Petrenko, V. A., Phage Display. *Chem Rev* **1997**, 97 (2), 391-410.
3. Esvelt, K. M.; Carlson, J. C.; Liu, D. R., A system for the continuous directed evolution of biomolecules. *Nature* **2011**, 472 (7344), 499-503.
4. Boder, E. T.; Wittrup, K. D., Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol* **2000**, 328, 430-44.
5. DePorter, S. M.; McNaughton, B. R., Engineered M13 bacteriophage nanocarriers for intracellular delivery of exogenous proteins to human prostate cancer cells. *Bioconjug Chem* **2014**, 25 (9), 1620-5.
6. DePorter, S. M.; Lui, I.; Bruce, V. J.; Gray, M. A.; Lopez-Islas, M.; McNaughton, B. R., Mutagenesis modulates the uptake efficiency, cell-selectivity, and functional enzyme delivery of a protein transduction domain. *Mol Biosyst* **2014**, 10 (1), 18-23.
7. DePorter, S. M.; Lui, I.; Mohan, U.; McNaughton, B. R., A protein transduction domain with cell uptake and selectivity profiles that are controlled by multivalency effects. *Chem Biol* **2013**, 20 (3), 434-44.
8. Torchilin, V., Intracellular delivery of protein and peptide therapeutics. *Drug Discov Today Technol* **2008**, 5 (2-3), e95-e103.
9. Fu, A.; Tang, R.; Hardie, J.; Farkas, M. E.; Rotello, V. M., Promises and pitfalls of intracellular delivery of proteins. *Bioconjug Chem* **2014**, 25 (9), 1602-8.

10. Stanzl, E. G.; Trantow, B. M.; Vargas, J. R.; Wender, P. A., Fifteen years of cell-penetrating, guanidinium-rich molecular transporters: basic science, research tools, and clinical applications. *Acc Chem Res* **2013**, *46* (12), 2944-54.
11. Fuchs, S. M.; Raines, R. T., Arginine grafting to endow cell permeability. *ACS Chem Biol* **2007**, *2* (3), 167-70.
12. Thompson, D. B.; Cronican, J. J.; Liu, D. R., Engineering and identifying supercharged proteins for macromolecule delivery into mammalian cells. *Methods Enzymol* **2012**, *503*, 293-319.
13. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R., A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. *Chem Biol* **2011**, *18* (7), 833-8.
14. Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* **2010**, *5* (8), 747-52.
15. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013**, *82*, 775-97.
16. Revets, H.; De Baetselier, P.; Muyldermans, S., Nanobodies as novel agents for cancer therapy. *Expert Opin Biol Ther* **2005**, *5* (1), 111-24.
17. Kubala, M. H.; Kovtun, O.; Alexandrov, K.; Collins, B. M., Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein Sci* **2010**, *19* (12), 2389-401.
18. Rothbauer, U.; Zolghadr, K.; Tillib, S.; Nowak, D.; Schermelleh, L.; Gahl, A.; Backmann, N.; Conrath, K.; Muyldermans, S.; Cardoso, M. C.; Leonhardt, H., Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods* **2006**, *3* (11), 887-9.

19. Pardon, E.; Laeremans, T.; Triest, S.; Rasmussen, S. G.; Wohlkonig, A.; Ruf, A.; Muyldermans, S.; Hol, W. G.; Kobilka, B. K.; Steyaert, J., A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc* **2014**, *9* (3), 674-93.
20. Abbady, A. Q.; Al-Mariri, A.; Zarkawi, M.; Al-Assad, A.; Muyldermans, S., Evaluation of a nanobody phage display library constructed from a Brucella-immunised camel. *Vet Immunol Immunopathol* **2011**, *142* (1-2), 49-56.
21. Sheikholeslami, F.; Rasaei, M. J.; Shokrgozar, M. A.; Dizaji, M. M.; Rahbarizadeh, F.; Ahmadvande, D., Isolation of a Novel Nanobody Against HER-2/neu Using Phage Displays Technology. *Labmedicine* **2010**, *41* (2), 69-76.
22. Saerens, D.; Pellis, M.; Loris, R.; Pardon, E.; Dumoulin, M.; Matagne, A.; Wyns, L.; Muyldermans, S.; Conrath, K., Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies. *J Mol Biol* **2005**, *352* (3), 597-607.
23. Conrath, K. E.; Lauwereys, M.; Galleni, M.; Matagne, A.; Frere, J. M.; Kinne, J.; Wyns, L.; Muyldermans, S., Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother* **2001**, *45* (10), 2807-12.
24. McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R., Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc Natl Acad Sci U S A* **2009**, *106* (15), 6111-6.
25. Guilleams, T.; El-Turk, F.; Buell, A. K.; O'Day, E. M.; Aprile, F. A.; Esbjorner, E. K.; Vendruscolo, M.; Cremades, N.; Pardon, E.; Wyns, L.; Welland, M. E.; Steyaert, J.; Christodoulou, J.; Dobson, C. M.; De Genst, E., Nanobodies raised against monomeric alpha-synuclein distinguish between fibrils at different maturation stages. *J Mol Biol* **2013**, *425* (14), 2397-411.

26. Appelbaum, J. S.; LaRochelle, J. R.; Smith, B. A.; Balkin, D. M.; Holub, J. M.; Schepartz, A., Arginine topology controls escape of minimally cationic proteins from early endosomes to the cytoplasm. *Chem Biol* **2012**, *19* (7), 819-30.
27. Rabideau, A. E.; Liao, X. L.; Pentelute, B. L., Delivery of mirror image polypeptides into cells. *Chem Sci* **2015**, *6* (1), 648-653.
28. Liao, X. L.; Rabideau, A. E.; Pentelute, B. L., Delivery of Antibody Mimics into Mammalian Cells via Anthrax Toxin Protective Antigen. *Chembiochem* **2014**, *15* (16), 2458-2466.

CHAPTER FOUR

Minimalist Antibodies and Mimetics: An Update and Recent Applications⁴

4.1 Introduction

Chapter One discusses the importance and thus great potential macromolecules including proteinaceous reagents have as clinical and basic research tools. One prime example includes antibodies. The immune system utilizes antibodies to recognize foreign or disease-relevant receptors, initiating an immune response to destroy unwelcomed guests. Because researchers can evolve antibodies to bind virtually any target, it is perhaps unsurprising that these reagents, and their small-molecule conjugates are used extensively in clinical and basic research environments. However, virtues of antibodies are countered by significant challenges. Foremost among these is the need for expression in mammalian cells (largely due to often necessary post translational modifications). In response to these challenges, researchers have developed an array of minimalist antibodies and mimetics, which are smaller, more stable, simpler to express in *Escherichia coli*, and are amenable to laboratory evolution and protein engineering. In this chapter, we describe these scaffolds and discuss recent applications of minimalist antibodies and mimetics.

⁴ Adapted from: Bruce, V.J.; Ta, A.N.; and McNaughton, B.R.; *ChemBioChem*, **2017**, 17, 1892. I co-authored this review article with Angeline Ta, a third-year graduate student at the time. Both of our individual projects deal with unique uses for minimalistic and stable antibody scaffold mimics giving us mutual knowledge and reason for collaborating on a review concerning the clinical and basic research applications for antibody derivatives and alternatives.

4.1.1 Antibodies: Structure, Function, Virtues, and Challenges

The most predominant antibody type is the immunoglobulin of isotype G (IgG), which weighs approximately 150 kDa.¹ Members of this antibody class each consist of two distinct regions: the fragment antigen-binding (Fab) and fragment crystallizable (Fc, **Figure 4.1A**). The Fab fragment consists of a constant light-chain domain (C_L) and a variable light-chain domain (V_L) (**Figure 4.1A**, green), linked to the constant ($CH1$) and the variable (V_H) heavy-chain domains (**Figure 4.1A**, gray). When folded properly, six solvent-exposed loops from V_L and V_H domains are presented. Collectively these loops are referred to as the complementary determining regions (CDRs, **Figure 4.1B**), and this is where the antigen is bound. Both the V_L and V_H domains display three CDRs, with loops having an average length of ten amino acids.

In contrast to the CDRs, the Fc has a greater sequence homology and common function. The Fcs of IgG1 and IgG4 each consist of two domains of the heavy chain ($CH2$ and $CH3$) connected to one another through two disulfide bonds in the hinge region (**Figure 4.1A**; all disulfide bonds are highlighted in red).² Multiple regions within the Fc are critical to antibody function, and endow unique properties. For example, the in vivo half-life of an IgG (ca. 21 days) is much longer than those of most proteins.³ This is achieved through an epitope on the surface of the Fc that interacts with the neonatal Fc receptor, FcRn. FcRn mediates a salvaging pathway by binding and transporting IgG into and across cells, dramatically slowing its degradation.⁴⁻⁵ FcRn binds to the $CH2$ - $CH3$ hinge region of IgG (**Figure 4.1C**) with high affinity under the acidic conditions typically found within endosomes ($pH < 6.5$) and with virtually no affinity in environments outside endosomes (typically $pH \approx 7.4$).⁶⁻⁷ This pH-dependent binding mediates the FcRn-IgG interaction after uptake into acidic endosomes, allowing IgG to piggyback with FcRn back to the plasma membrane and the complex to dissociate once returned to the circulatory system. This

sequestration/transport mechanism saves antibodies from degradation through the endosome/lysosome pathway. Regions of Fc also mediate immune system stimulation, such as antibody-dependent cellular cytotoxicity (ADCC)⁸⁻⁹ or complement-dependent cytotoxicity (CDC)¹⁰⁻¹². Fc gamma receptors (FcγRs) on the surfaces of immune effector cells such as natural killer (NK) cells and macrophages recognize the Fc regions of antibodies bound to a target cell (**Figure 4.1D**).¹³⁻¹⁵ Upon binding, the immunoreceptor tyrosine-based activation motif (ITAM) is phosphorylated; this then triggers the activation of the effector cell and release of perforin, lytic enzymes, tumor necrosis factor (TNF), and/or granzymes for cell destruction through ADCC. In CDC, C1q of complex C1 binds to the Fc region and triggers the complement cascade activation and eventual formation of membrane attack complex (MAC) at the surfaces of the target cells, leading to cell lysis.

In addition to the proteinogenic amino acids that make up antibodies, extensive post-translational modifications (disulfide-bond formation and glycosylation) are required to deliver a mature immunoglobulin. As a result of the large size and molecular complexity of an antibody, challenges arise in their production – fully modified antibodies must largely be prepared in mammalian cells.¹⁶⁻¹⁷ This form of production is relatively costly, slow, and low-yielding in comparison with expression of many recombinant proteins in *E. coli*.

It is these challenges in the preparation and manipulation of antibodies that has motivated researchers to develop minimalist forms and mimetics with improved expression in *E. coli* and stability. Over the past few decades, researchers have developed a number of protein scaffolds that

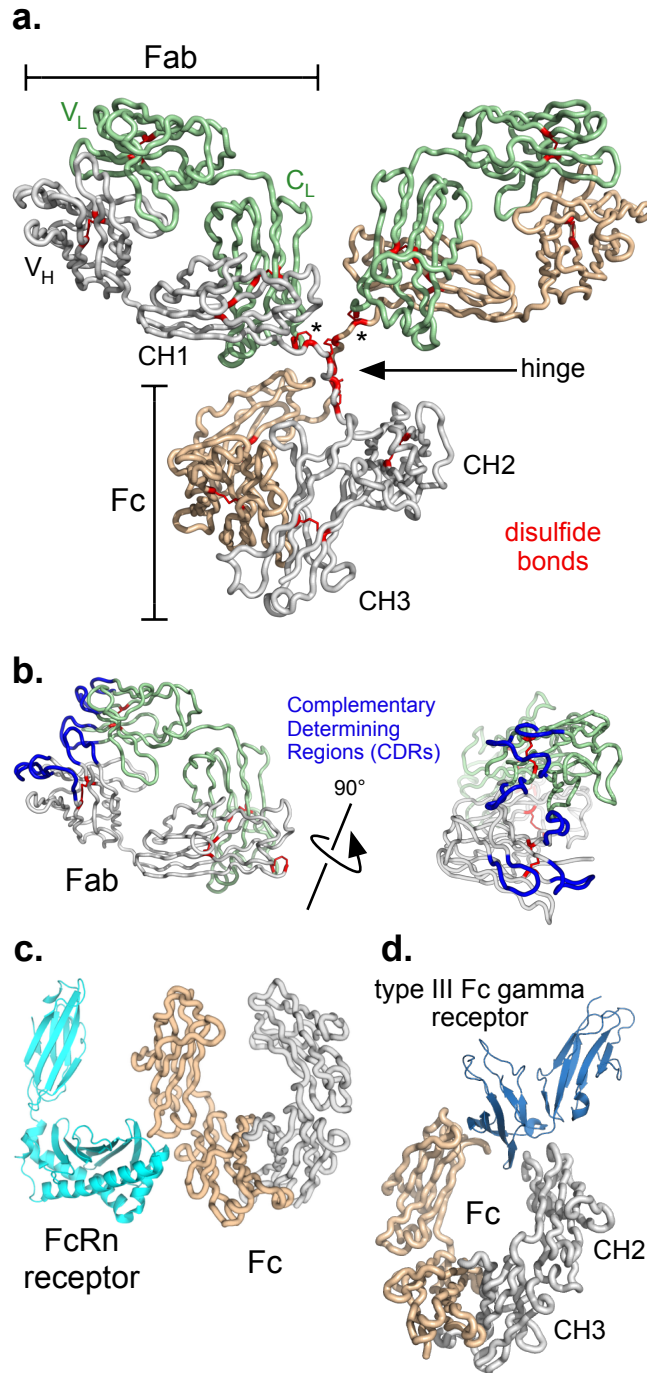


Figure 4. 1 Typical Antibody Structure. (a) Immunoglobulin isotype G (IgG) consists of two heavy chains (grey and brown) and two light chains (green). The heavy chains contain the fragment crystallizable (Fc), the constant region (CH1) and variable region (V_H). The light chain is made up of a constant region (C_L) and variable region (V_L), and is covalently attached to V_H/C_H by a single disulfide bond (highlighted by *). Collectively, the V_L/C_L-CH1/V_H region is called the Fab fragment, and is where antigen binding occurs (antigen receptors are highlighted in blue, PDB: 1IGY). (b) Complementary determining regions (CDRs, blue) from a Fab fragment. (c) Interaction between CH3 and CH2 of Fc and the neonatal Fc receptor (FcRn), (PDB: 1I1A). (d) ADCC is initiated by the type III Fc gamma receptor binding to CH3 domains from the A and B chain of Fc (PDB: 1T89).

are amenable to extensive mutagenesis and laboratory evolution to achieve new recognition and unique function with relative ease.

Many new protein-protein interactions can be achieved and characterized through high-throughput screening or in laboratory evolution by resurfacing helix or β -strand structural features.¹⁸⁻²³ In this chapter, however we focus on minimalist forms of antibodies, and their mimetics, which, like antibodies, rely on maturation of loops to achieve recognition. Modern applications and protein engineering efforts to generate new properties and function are discussed throughout.

4.2 Antibody Fragments

Immunoglobulins, including the isotype IgG, have a modular architecture, and each module has a unique biological function. In particular, it is this modular architecture that allows researchers to minimize components, thus generating new proteins that retain certain desired properties, but intelligently lose others. For example, because the Fab fragment is solely responsible for antigen recognition and binding, this domain – and variations on that structural theme – have been used for recognition in clinical and basic research settings, as well as for targeted delivery of cargo.

4.2.1 Fab Fragments

Cartoon depictions of IgG and antibody fragments discussed in this chapter are shown in **Figure 4.2**. Full-length Fab fragments contain both CH1/V_H and C_L/V_L fusions, connected by single disulfide bonds (denoted by asterisks in **Figure 4.1**). These minimalist antibodies have some advantages and disadvantages of immunoglobulins. For example, Fab fragments retain target recognition, but lose properties encoded within the Fc domain, such as immune response

stimulation and extended *in vivo* half-life. Because these fragments have a relatively short existence in plasma, in comparison with full-length antibodies, antibody fragments might be of particular value in applications that favor or require shorter biological lifetime (such as imaging). Moreover, their small size allows deeper tumor penetration, simpler expression in *E. coli*, and manipulation in the laboratory.²⁴⁻²⁶

Fab fragments can be produced through chemical or protease digestion of full-length immunoglobulins.²⁷ Additionally, these fragments can be produced by recombinant expression in bacteria.²⁸ Fab fragments have been used as therapeutics, as well as in diagnosis, detection, imaging, and crystallography applications.^{24, 27, 29-30}

At present, a number of Fab fragments are in clinical trials. For example, citatuzumab bogatox (VB6-845) is a recombinant immunotoxin for use as a treatment for ovarian cancer and other solid tumors.³¹ In this drug lead, deBouganin – a de-immunized plant toxin – is fused to a humanized Fab fragment that targets epithelial cell adhesion molecules (EpCAMs). In addition, naptumomab estafenato (ABR-217620) is a fusion protein therapeutic for advanced renal cell carcinoma and other solid tumors.³² The fusion consists of a Fab fragment that binds 5T4 (a cell-surface tumor antigen), and superantigen staphylococcal enterotoxin A (a protein that binds to major histocompatibility complex class II molecules and activates T lymphocytes).

Fab fragments are also used as tools for imaging and detection. ThromboView is a radiolabeled Fab fragment that targets the D dimer region of crosslinked fibrin for deep-vein thrombosis imaging.³³ In basic research applications, Fab fragments have been used for imaging inside mammalian cells. For example, Stasevich, Morisaki, and co-workers have utilized Fabs to study translation in living cells. Using a Fab fragment that recognizes the FLAG tag (DYKDDDK) they perform nascent chain tracking (NCT).³⁴ mRNA encoding a 10X FLAG-tagged protein and 24X

MS2 tag in the 3'-untranslated region is produced in cells. Fluorescently labeled MS2 coat protein recognizes the mRNA, thus detecting its presence in a mammalian cell. By use of an orthogonally labeled FLAG binding Fab, translation of that encoded protein is detected, following translation of the FLAG tag from the ribosome. Collectively, these two fluorescently labeled components provide a glimpse into translational dynamics in living mammalian cells.

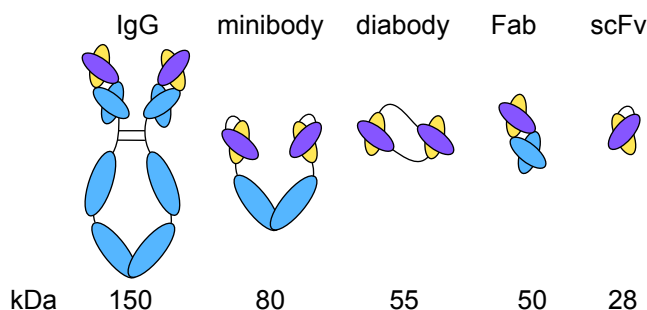


Figure 4. 2 Cartoon depiction of IgG and fragments discussed in this Chapter.

4.2.2 Single-chain Fragment Variables (scFvs)

Single-chain fragment variable antibodies (scFvs) were first reported in 1988 as minimalist forms of Fab fragments.³⁵ These ≈ 28 kDa fragments result from the genetic linkage of V_H to V_C typically with a flexible 10-to-25-residue linker.²⁴ Whereas antibodies can contain up to 25 disulfide bonds and Fab fragments can require five disulfide bonds, scFv typically only contain two, thus simplifying their recombinant expression and stability in reducing environments (such as the cytosol of bacteria). New strains of *E. coli* with enzymes to facilitate disulfide formation further simplify the recombinant production of scFvs.³⁶

Similarly to other Fab-based antibody fragments, scFvs do not participate in immune response stimulation, and removal of the FcRn receptor results in substantially decreased *in vivo* half-lives. However, their small size and easy expression make scFvs relatively simple minimalist antibodies to prepare and manipulate in the laboratory. Clinically, scFvs display better tumor penetration,

more rapid blood clearance, lower retention times in non-target tissue, and reduced immunogenicity.³⁷

Because functional scFvs can be expressed in the reducing environment of the cytosol, these reagents can be generated inside a cell for use in certain applications. For example, Kimura, Sato, and co-workers utilized a scFv specific for histone H3 lysine 9 acetylation (H3K9ac), fused with GFP to identify post-translational histone modifications in living cells.³⁸ This approach enables tracking of the spatiotemporal dynamics of endogenous histone modification in a genetically encoded format.

Recently, scFv-based technology called SunTag has been used for real-time detection of proteins in living cells and to amplify transcription.³⁹ In the context of protein detection, cells are made to express a protein that displays many copies of a short peptide epitope. The cells also express a scFv that recognizes the epitope, fused to a fluorescent protein such as GFP. When concomitantly expressed, the scFv-GFP fusion selectively recognizes the epitope-tagged protein, resulting in the illumination of that protein in a living cell (**Figure 4.3A**). Additionally, SunTag has been used to enhance transcription. Vale, Tanenbaum, and co-workers made cells express a nuclease-inactive form of Cas9 (dCas9) fused to multiple peptide epitopes, as well as a scFv that binds the epitope while fused to VP64, a transcriptional activator. Complex formation between the peptide epitopes on dCas9 and the scFv-VP64 fusion led to recruitment of many copies of the transcriptional activator to transcriptional machinery on DNA – resulting in increased transcription (**Figure 4.3B**).

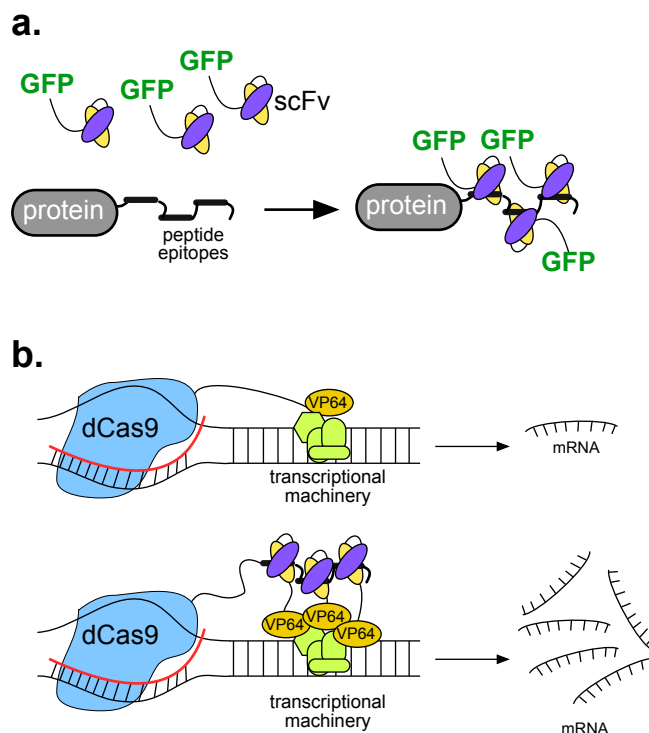


Figure 4. 3 Utility of scFv fusions in SunTag Technology. (A) to illuminate and track proteins in living cells, and (B) to recruit a transcriptional activator (VP64), resulting in increased transcription of a gene in a cell.

4.2.3 Minibodies

A minibody is a single polypeptide consisting of scFv-CH3-CH3-scFv and can be expressed as a functional protein through recombinant expression in *E. coli*.⁴⁰ The principal benefit of including the CH3 is an appreciable increase in biological half-life relative to scFv. Variants that contain the hinge region (flex minibodies) and variants that do not contain the hinge region (LD minibodies) have both been reported – with the flex minibodies showing higher tumor uptake and slower clearance times.⁴⁰ Similarly to their scFv cousins, minibodies retain target affinity, but lose immune response stimulation.

However, conjugates of minibodies have been used for targeted delivery of toxic proteins or small-molecule compounds.⁴⁰ Additionally, because a minibody consists of a single polypeptide,

scFv domains with different target recognition can be encoded, and bispecific binding (concomitant recognition of two different targets) can be achieved.⁴¹

Recently, the Wu lab developed an immunoPET (positron emission tomography) radiotracer for imaging of prostate cancer by targeting prostate stem cell antigens through affinity maturation of a previously developed hu1G8 minibody modified with ¹²⁴I and ⁹⁸Zr radiolabels.⁴²⁻⁴³ Marasco, Han, and co-workers reported a minibody as a potential therapy for cutaneous T-cell lymphoma (CTCL) that targets the CC chemokine receptor 4 (CCR4).⁴⁴ In a creative modern application, Park, Lee, and co-workers prepared a polypeptide consisting of a polyarginine (Arg₉) and an anti-JL1 minibody. When this was noncovalently complexed with siRNA, through charge complementation, and then applied to mammalian cells, siRNA delivery was achieved specifically in leukemic cells.⁴⁵

In another innovative application, Marasco, Abdel-Motal, and co-workers examined the utility of an anti-gp120 minibody in protecting against sexual transmission of HIV-1.⁴⁶ Through the use of an adeno-associated viral (AAV) vector, anti-HIV-1 gp120 minibody was introduced into cervico-vaginal epithelial cells. After secretion to the cell surface, the minibody binds HIV-1 gp120, resulting in sequestration of the virus and decreased infection.

4.2.4 Diabodies

A diabody is a complex consisting of two unconjugated single-chain fragment variables.⁴⁷⁻⁴⁸ Although scFvs can be engineered to be multivalent with the addition of either chemical or genetic cross-links, it was found that reducing the scFv linker allowed for multimerization and stability. As in the case of minibodies, because scFv domains with differing targets can be mixed, bispecific recognition can be achieved.

Similarly to their larger minibody relatives, diabodies have relatively short *in vivo* half-lives, and thus could be better suited for imaging, because they can illuminate their targets and then be degraded and cleared. Additionally, the smaller size of diabodies, in relation to IgG and larger fragments, endows improved accumulation and penetration of tumors expressing relevant antigens.⁴⁹⁻⁵⁰ With these characteristics in mind, most therapeutically relevant applications of diabodies have revolved around PET imaging. At present, diabody conjugates to PET labels have been validated for pancreatic cancer (anti-CAI9-9), anti-leukocyte cell-adhesion molecule (ALCAM/CD166), and breast cancer (anti-HER2).⁵¹⁻⁵³

Diabodies conjugates for imaging applications that involve binding of the extra domain-B (EDB) of fibronectin (a biomarker for angiogenesis/atherosclerotic plaque), carcinoembryonic antigen (CEA, a validated marker for gastrointestinal cancers), and Her2/Neu (a biomarker for both ovarian and breast cancers) have also been reported.⁵⁴⁻⁵⁶

4.3 Nanobodies – A Camelid-Derived Scaffold

Heavy-chain IgGs (hcIgGs) produced in camelids differ from IgGs produced in other mammals.⁵⁷⁻⁵⁸ Although hcIgG is also a homodimer of two disulfide-linked heavy chains, with familiar CH2, CH3, and variable domains, it lacks a light chain. Binding between antigen and hcIgG, as for their IgG cousins, relies entirely on amino acids residing in loops (complementarity-determining regions, CDRs) of the single variable domain (referred to as V_HH in hcIgG, **Figure 4.4A**). Separating the CDR loops are four relatively sequence-homologous β -strands, which makeup the “framework region”.⁵⁷ When separated from hcIgG, the V_HH domain is called a “nanobody” (**Figure 4.4B**), and an excellent review on their discovery and structure has been published.⁵⁹

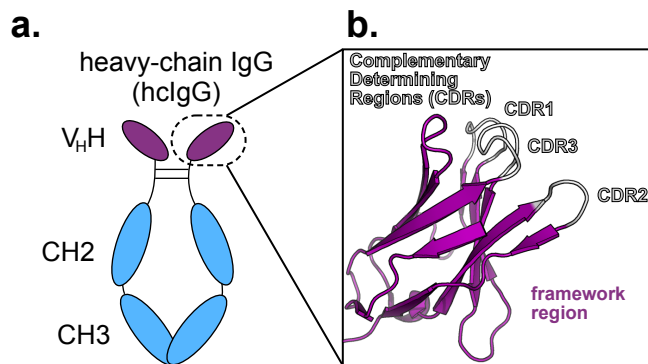


Figure 4. 4 The Architecture of hIgG and Nanobodies. (A) Heavy-chain IgG (hclgG), which consists of two heavy chains (CH3, CH2, and V_HH) connected by disulfide bonds in the hinge region. (B) A nanobody – GFP binding nanobody is shown as an example (PDB: 3OGO) – is the V_HH domain of a hclgG and consists of a framework region (purple) and CDRs (gray) where antigen recognition occurs.

As discussed in Chapter Three, nanobodies have many properties that make them particularly well suited as scaffolds for the directed evolution of new recognition in the laboratory.⁶⁰ These proteins are small (≈ 15 kDa), can be expressed in a folded and stable form with or without disulfide bonds in *E. coli*, and are easy to manipulate in the laboratory. An obvious difference between the variable regions of IgG and hclgG is that binding is generated from amino acids in three loops in the hclgG variant and in six loops in IgG (from heavy and light variable domains). However, hclgG can compensate for this smaller apparent binding surface by expansion of CDR loops. For example, in comparison with IgG, nanobodies typically have longer CDR3 loops (ranging from eight to 24 residues) than those found in mouse or human antibodies (nine and 12 residues, respectively).⁶¹ The expanded CDR3 can dramatically increase the size of the paratope (the part of the protein that recognizes the epitope). This extended display architecture is generally credited with allowing nanobodies to bind surfaces that challenge or evade IgG, such as deep clefts within enzymes.⁶¹⁻⁶²

4.3.1 Recent applications of nanobodies

Like their Fab fragment counterparts, nanobodies do not contain FcRn receptors and thus have relatively short in vivo half-lives. As a result, nanobodies can be used in situations such as bioimaging, in which relatively short half-lives and clearances are favored. Of course, this requires selective recognition of a disease-relevant cell surface biomarker. Probably one of the most widely studied and utilized biomarkers is the human epidermal growth factor receptor type 2 (HER2), which is overexpressed in $\approx 20\text{-}30\%$ of breast and ovarian carcinomas.⁶³⁻⁶⁴ This tyrosine kinase receptor is responsible for cell proliferation, reduction in apoptosis, and enhanced cell mobility, making it an ideal extracellular model protein. A nanobody for the HER2 receptors has been developed (termed 5F7, $K_D \approx 0.1$ nM) and used extensively for imaging and proof-of-concept nanobody technologies.⁶⁵⁻⁶⁶

With no receptors to recruit NK cells or other immune system components present in the Fc region, nanobodies cannot illicit immune responses such as ADCC or CDC. One approach would be to fuse a nanobody to a Fc dimer. However, post-translational glycosylation of Fc is necessary to induce ADCC or CDC, and this requires expression in mammalian cells, thus complicating its preparation in the lab. To overcome this obstacle, our lab has prepared conjugates of small-molecule compounds and nanobodies that bind a target cell biomarker and recruit an antibody to the cell surface, resulting in ADCC (**Figure 4.5A**).⁶⁷ Specifically, using a combination of lipoic acid ligase bioconjugation and reactivity between a hydrazine and a protein-bound aldehyde, we coupled dinitrophenyl (DNP) to a previously reported HER2-binding nanobody.⁶⁸ As a result of human exposure to DNP, likely from DNP-containing dyes, preservatives, and/or pesticides, it is estimated that $\approx 1\%$ of IgGs and IgMs recognize DNP,⁶⁹ and can thus recruit endogenous antibodies to a targeted cell.⁷⁰⁻⁷¹ Satisfyingly, when HER2-positive breast cancer cells (SK-BR-3)

were treated with the nanobody activation immunotherapeutic, anti-DNP antibody and peripheral blood mononuclear cells (PBMCs) were recruited, triggering appreciable ADCC. Conversely, the nanobody activation immunotherapeutic did not induce ADCC for MB-MDA-231 breast cells that express low levels of HER2. Similarly, when either cell line was treated under conditions in which a component (the nanobody activation immunotherapeutic, anti-DNP antibody, or PBMC) was absent, no appreciable ADCC was detected (**Figure 4.5B**).⁶⁷

Chapter One discusses a huge limitation for most proteins, including antibodies, their fragments, and mimetics, in their inability to appreciably penetrate mammalian cells; nanobodies are no exception. This limits their recognition to cell-surface or excreted proteins. However, because of the robust nature of nanobodies, several intracellularly active nanobodies have been identified and remain functional in reducing environments, such as the interior of a mammalian cell.⁷² Chapter Three discusses our efforts to prepare nanobodies that actively penetrate mammalian cells, through polycationic resurfacing (mutation of solvent-exposed residues either to lysine or to arginine) on three separate nanobody frameworks (a resurfaced GFP-binding nanobody⁷³⁻⁷⁴ is shown as an example in **Figure 4.5C**).⁷⁵

Whereas nanobodies have largely been used to recognize large conformational regions on proteins, they can also be subjected to evolution to recognize small unfolded peptide epitopes. Recently, Rothbauer, Braun, and co-workers generated a nanobody called BC2 that recognizes a short linear epitope corresponding to residues 16-27 of β -catenin (BC2T).⁷⁶ The structure of this complex was solved by X-ray crystallography, revealing complete encapsulation of the epitope by an extended CDR3 loop (**Figure 4.5D**). This nanobody has shown utility as a reagent for capture and detection of BC2T-tagged proteins. Chapter Five discuss effort by our laboratory to expand this nanobodies utility and evaluated it as protein fusions and bioconjugates to fluorophores and

enzymes as alternative bioanalytical reagents to antibodies in the capacity of enzyme-linked immunosorbent assays (ELISAs), flow cytometry, and Western blot analysis.⁷⁷ This is the subject of Chapter Five.

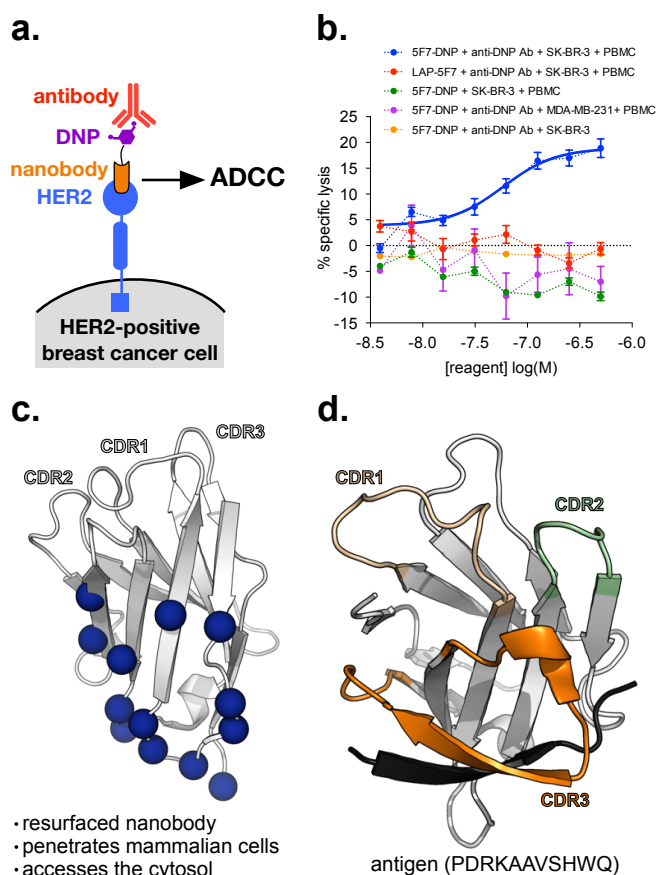


Figure 4.5 Recent Applications of Nanobodies. (A) Concept of a nanobody activation immunotherapeutic. A HER2-binding nanobody (orange) is chemically conjugated to dinitrophenyl (DNP, purple), which is recognized by endogenous antibodies in human serum (red). Recruitment of antibodies to the surface of HER2-positive breast cancer cells leads to antibody-dependent cellular cytotoxicity (ADCC). (B) ADCC of high HER2-expressing SK-BR-3 cells (blue), but not of low HER2-expressing MB-MDA-231 cells (purple), triggered by the nanobody activation immunotherapeutic DNP-5F7. The unconjugated nanobody (LAP-5F7) is not toxic to SK-BR-3 cells (red). DNP-5F7 in the absence of anti-DNP antibodies (green) or PMBCs (orange) also does not trigger ADCC. Error bars represent standard error from three independent experiments. (C) Polycationic resurfacing of the nanobody framework region results in potent cell penetration and access to the cytosol. Residues highlighted with a blue sphere were mutated to either arginine or lysine (PBD: 3OGO). (d) Structure of a recently reported nanobody with an expanded CDR3 that is structured and able to bind a small peptide antigen (PBD: 5IVO).

4.4 Monobodies – A Fibronectin-Derived Scaffold

All of the above examples in this Chapter (variations on the theme of Fab fragments and nanobodies) are derived from immunoglobulins (either IgGs or hcIgGs). These are contrasted with monobodies – scaffolds from the human derived 10th fibronectin type III domain (FNfn10). With the aid of the FNfn10 scaffold, protein binding interactions can be fashioned to particular targets through loop interactions or side-and-loop interactions. Loops can be mutated and elongated with minimal stability loss, allowing for a large diversity of binding faces. Because they do not start from the usual protein scaffolds involved in adaptive immunity, monobodies can be made to bind to a variety of targets to serve many different functions while still being inherently nontoxic and immunogenic. First reported by the Koide lab in 1998, monobodies are essentially structurally simplified mimics of a heavy-chain fragment variable, in that both present three binding loops for antigen recognition (**Figure 4.6**).⁷⁸ Monobodies do not contain disulfide bonds, are small (≈ 10 kDa) and generally stable, express well as soluble proteins in *E. coli*, and, due to the nature of the fibronectin type III structure from which monobodies are derived can be used as a binding protein that mimics IgG V_H (**Figure 4.6 A-C**). Similarly to Fab fragments and nanobodies, monobodies are useful in binding to a specific target, but do not contain an Fc region, therefore dramatically decreasing their serum stability in comparison to antibodies. This results in monobodies generally being used as diagnostic tools to identify cell-surface biomarkers (where relatively quick clearance might be beneficial), and more recently as modulators of enzyme function and selectivity.

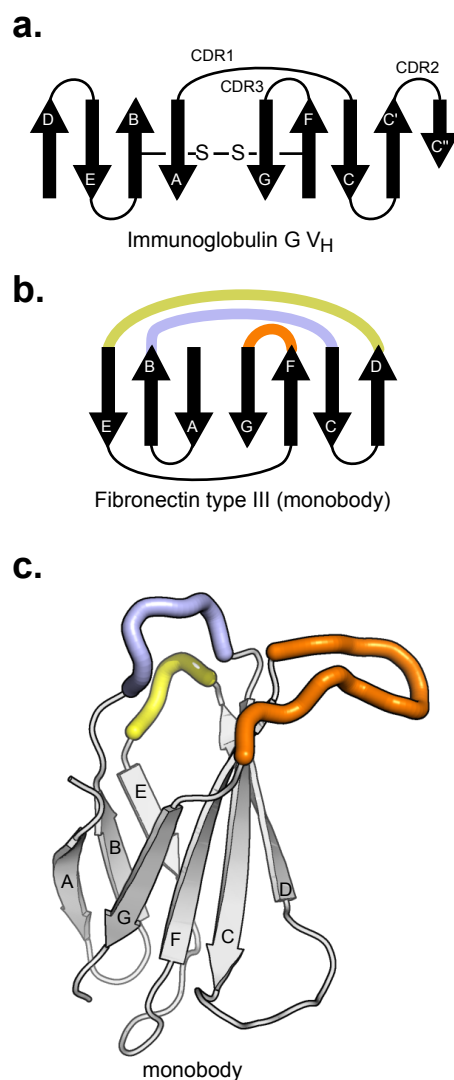


Figure 4.6 Structural comparison of IgG and Monobodies. (A) Structural architecture of IgG V_H. (B) Structural architecture of the tenth fibronectin type III (monobody). (C) Structure of a monobody. Antigen binding loops are colored (PBD: 1FNF).

4.4.1 Recent applications of monobodies

Monobodies have been used as proteinaceous reagents to bind various disease-relevant macromolecules, resulting in the modulation, study, and characterization of complex cellular processes. For example, researchers have used monobodies that bind Fluc-type F⁻ channels to validate its unique mechanism of action for controlling intracellular levels of fluoride ion.⁷⁹⁻⁸¹ Monobody drug leads that bind structurally diverse disease-relevant targets have also been

reported. For example, Kuhlman, Guntas, and co-workers used computational loop grafting of the BC and FG loops, together with phage display to engineer a monobody (R1) that binds to Kelch-like ECH-associated protein 1 (KEAP1) with a K_D of 300 pM. This monobody inhibits the interaction between KEAP1 and nuclear factor erythroid 2-related factor 2 (NRF2), resulting in activation of NRF2⁸² – a key regulator of cellular oxidative environments and an interaction associated with several disease states.

Monobodies have also been used as diagnostic reagents. For example, Hong, Park, and co-workers developed a monobody that binds to human EphA2 (hEphA2), an early marker for various tumors.⁸³ Whereas monobodies have largely been used to recognize and/or modulate the biological activation of specific proteins, they have more recently been applied to more diverse functions such as altered enzyme activity and biotechnology validation. In recent work, Koide, Tanaka, and co-workers showed that a monobody can alter an enzyme's specificity for its target, without modifying the amino acid sequence of the enzyme.⁸⁴ In particular, they found monobodies that were able to restrict β -galactosidase transgalactosylation yields of galacto-oligosaccharides (GOSs) to specific lengths rather than mixtures.

Very recently, in the context of extending phage-assisted continuous evolution (PACE)⁸⁵ to protein-protein interaction discovery, Liu, Badran, and co-workers re-evolved a monobody to bind the SH3 domain of ABL1. Beginning with a previously characterized mutant monobody (Tyr87Ala), which binds the SH2 domain target with dramatically lower affinity (100 -1000 – fold), continuous evolution through PACE was used to regain tight binding.⁸⁶

4.5 Conclusion

Throughout this thesis, a common theme is the increasing use of proteins in basic research and clinical applications to modulate disease-relevant receptors and to control cell function and fate. At present, seven of the ten top selling drugs are biologics, and many of these are antibodies and their conjugates. Relative challenges associated with the expression of full-length and chemically mature (post-translationally modified) antibodies in mammalian cell culture have motivated researchers, including us (Chapter Five and Six) to develop an array of minimalist antibody forms and mimetics that still experience potent recognition of epitope. In this chapter, we have highlighted various forms of IgG fragments (Fab, scFv, minibodies, diabodies), hIgG-derived nanobodies, and fibronectin-derived monobodies as alternatives to full-length IgGs.

Specifically, scFvs, nanobodies, and monobodies are structurally simpler (lack disulfide bonds), easier to express in *E. coli*, and can be simpler to engineer and use in directed evolution than IgGs and IgG-derived counterparts. Moreover, the relatively small sizes of scFvs, nanobodies, and monobodies often correlate with greater tumor penetration, and thus, in some cases, this virtue could be used to improve the efficacy of tumor-targeted therapies (although challenges with relatively short *in vivo* lifetimes would remain an issue). Because long *in vivo* lifetimes are a key advantage of IgGs and derivatives that retain Fc, many applications of IgG and derivatives that retain, many applications of IgG derivatives lacking Fc, or of non-immunoglobulin-derived proteins such as nanobodies and monobodies, include diagnostic and bioanalytical applications. However, smaller IgG-derived proteins that lack Fc, as well as nanobodies and monobodies, are particularly well suited for some medically relevant applications, such as bioimaging, because rapid clearance is not a major issue, or even beneficial. Historically, full-length antibodies as bioanalytical tools – Western blot being an obvious example. However, relative challenges in the

expression of these molecular Winnebagos opens the door for smaller and simpler proteins, and nanobodies have recently been used in this context as well as other assays (discussed in Chapter Five).⁷⁶⁻⁷⁷

Full-length antibodies continue to enjoy application in immunotherapy and as conjugates to small-molecule therapeutics and imaging reagents, in which their primary job is to deliver these cargos selectively to diseased cells.⁸⁷ Owing to their simpler expression and ability for straightforward evolution in the laboratory, truncated structural forms of antibodies, and non-immunoglobulin mimetics play an increasingly important role in human health. We believe the creative use of these proteins will continue to represent a growing area of protein science, biologics, research, and therapeutic discovery.

REFERENCES

1. Vidarsson, G.; Dekkers, G.; Rispens, T., IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* **2014**, *5*, 520.
2. Liu, H.; May, K., Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs* **2012**, *4* (1), 17-23.
3. Suzuki, T.; Ishii-Watabe, A.; Tada, M.; Kobayashi, T.; Kanayasu-Toyoda, T.; Kawanishi, T.; Yamaguchi, T., Importance of neonatal FcR in regulating the serum half-life of therapeutic proteins containing the Fc domain of human IgG1: a comparative study of the affinity of monoclonal antibodies and Fc-fusion proteins to human neonatal FcR. *J Immunol* **2010**, *184* (4), 1968-76.
4. Goebel, N. A.; Babbey, C. M.; Datta-Mannan, A.; Witcher, D. R.; Wroblewski, V. J.; Dunn, K. W., Neonatal Fc receptor mediates internalization of Fc in transfected human endothelial cells. *Mol Biol Cell* **2008**, *19* (12), 5490-505.
5. Roopenian, D. C.; Akilesh, S., FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* **2007**, *7* (9), 715-25.
6. Burmeister, W. P.; Gastinel, L. N.; Simister, N. E.; Blum, M. L.; Bjorkman, P. J., Crystal structure at 2.2 Å resolution of the MHC-related neonatal Fc receptor. *Nature* **1994**, *372* (6504), 336-43.
7. Burmeister, W. P.; Huber, A. H.; Bjorkman, P. J., Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* **1994**, *372* (6504), 379-83.

8. Ravetch, J. V.; Clynes, R. A., Divergent roles for Fc receptors and complement in vivo. *Annu Rev Immunol* **1998**, *16*, 421-32.
9. Seidel, U. J.; Schlegel, P.; Lang, P., Natural killer cell mediated antibody-dependent cellular cytotoxicity in tumor immunotherapy with therapeutic antibodies. *Front Immunol* **2013**, *4*, 76.
10. Prang, N.; Preithner, S.; Brischwein, K.; Goster, P.; Woppel, A.; Muller, J.; Steiger, C.; Peters, M.; Baeuerle, P. A.; da Silva, A. J., Cellular and complement-dependent cytotoxicity of Ep-CAM-specific monoclonal antibody MT201 against breast cancer cell lines. *Br J Cancer* **2005**, *92* (2), 342-9.
11. Walport, M. J., Complement. Second of two parts. *N Engl J Med* **2001**, *344* (15), 1140-4.
12. Walport, M. J., Complement. First of two parts. *N Engl J Med* **2001**, *344* (14), 1058-66.
13. Radaev, S.; Motyka, S.; Fridman, W. H.; Sautes-Fridman, C.; Sun, P. D., The structure of a human type III Fcgamma receptor in complex with Fc. *J Biol Chem* **2001**, *276* (19), 16469-77.
14. Nimmerjahn, F.; Ravetch, J. V., Fcgamma receptors as regulators of immune responses. *Nat Rev Immunol* **2008**, *8* (1), 34-47.
15. Nimmerjahn, F.; Ravetch, J. V., Fc-receptors as regulators of immunity. *Adv Immunol* **2007**, *96*, 179-204.
16. Sapra, P.; Shor, B., Monoclonal antibody-based therapies in cancer: advances and challenges. *Pharmacol Ther* **2013**, *138* (3), 452-69.
17. Samaranayake, H.; Wirth, T.; Schenkwein, D.; Raty, J. K.; Yla-Herttuala, S., Challenges in monoclonal antibody-based therapies. *Ann Med* **2009**, *41* (5), 322-31.
18. Lofblom, J.; Frejd, F. Y.; Stahl, S., Non-immunoglobulin based protein scaffolds. *Curr Opin Biotechnol* **2011**, *22* (6), 843-8.

19. Silverman, A. P.; Kariolis, M. S.; Cochran, J. R., Cystine-knot peptides engineered with specificities for $\alpha(\text{IIb})\beta(3)$ or $\alpha(\text{IIb})\beta(3)$ and $\alpha(\text{v})\beta(3)$ integrins are potent inhibitors of platelet aggregation. *J Mol Recognit* **2011**, *24* (1), 127-35.
20. Gebauer, M.; Skerra, A., Engineered protein scaffolds as next-generation antibody therapeutics. *Curr Opin Chem Biol* **2009**, *13* (3), 245-55.
21. Skrlec, K.; Strukelj, B.; Berlec, A., Non-immunoglobulin scaffolds: a focus on their targets. *Trends Biotechnol* **2015**, *33* (7), 408-18.
22. Binz, H. K.; Amstutz, P.; Pluckthun, A., Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol* **2005**, *23* (10), 1257-68.
23. Chapman, A. M.; McNaughton, B. R., Scratching the Surface: Resurfacing Proteins to Endow New Properties and Function. *Cell Chem Biol* **2016**, *23* (5), 543-53.
24. Holliger, P.; Hudson, P. J., Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* **2005**, *23* (9), 1126-36.
25. Yokota, T.; Milenic, D. E.; Whitlow, M.; Schlom, J., Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms. *Cancer Res* **1992**, *52* (12), 3402-8.
26. Jain, R. K., Physiological barriers to delivery of monoclonal antibodies and other macromolecules in tumors. *Cancer Res* **1990**, *50* (3 Suppl), 814s-819s.
27. Rader, C., Overview on concepts and applications of Fab antibody fragments. *Curr Protoc Protein Sci* **2009**, Chapter 6, Unit 6 9.
28. Better, M.; Chang, C. P.; Robinson, R. R.; Horwitz, A. H., Escherichia coli secretion of an active chimeric antibody fragment. *Science* **1988**, *240* (4855), 1041-3.
29. Nelson, A. L.; Reichert, J. M., Development trends for therapeutic antibody fragments. *Nat Biotechnol* **2009**, *27* (4), 331-7.

30. Nelson, A. L., Antibody fragments: hope and hype. *MAbs* **2010**, 2 (1), 77-83.
31. Entwistle, J.; Brown, J. G.; Chooniedass, S.; Cizeau, J.; MacDonald, G. C., Preclinical evaluation of VB6-845: an anti-EpCAM immunotoxin with reduced immunogenic potential. *Cancer Biother Radiopharm* **2012**, 27 (9), 582-92.
32. Eisen, T.; Hedlund, G.; Forsberg, G.; Hawkins, R., Naptumomab estafenatox: targeted immunotherapy with a novel immunotoxin. *Curr Oncol Rep* **2014**, 16 (2), 370.
33. Douketis, J. D.; Ginsberg, J. S.; Haley, S.; Julian, J.; Dwyer, M.; Levine, M.; Eisenberg, P. R.; Smart, R.; Tsui, W.; White, R. H.; Morris, T. A.; Kaatz, S.; Comp, P. C.; Crowther, M. A.; Kearon, C.; Kassis, J.; Bates, S. M.; Schulman, S.; Desjardins, L.; Taillefer, R.; Begelman, S. M.; Gerometta, M., Accuracy and safety of (99m)Tc-labeled anti-D-dimer (DI-80B3) Fab' fragments (ThromboView(R)) in the diagnosis of deep vein thrombosis: a phase II study. *Thromb Res* **2012**, 130 (3), 381-9.
34. Morisaki, T.; Lyon, K.; DeLuca, K. F.; DeLuca, J. G.; English, B. P.; Zhang, Z.; Lavis, L. D.; Grimm, J. B.; Viswanathan, S.; Looger, L. L.; Lionnet, T.; Stasevich, T. J., Real-time quantification of single RNA translation dynamics in living cells. *Science* **2016**, 352 (6292), 1425-9.
35. Huston, J. S.; Levinson, D.; Mudgett-Hunter, M.; Tai, M. S.; Novotny, J.; Margolies, M. N.; Ridge, R. J.; Bruccoleri, R. E.; Haber, E.; Crea, R.; et al., Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc Natl Acad Sci U S A* **1988**, 85 (16), 5879-83.
36. Matos, C. F.; Robinson, C.; Alanen, H. I.; Prus, P.; Uchida, Y.; Ruddock, L. W.; Freedman, R. B.; Keshavarz-Moore, E., Efficient export of prefolded, disulfide-bonded recombinant proteins

- to the periplasm by the Tat pathway in *Escherichia coli* CyDisCo strains. *Biotechnol Prog* **2014**, *30* (2), 281-90.
37. Weisser, N. E.; Hall, J. C., Applications of single-chain variable fragment antibodies in therapeutics and diagnostics. *Biotechnol Adv* **2009**, *27* (4), 502-20.
38. Sato, Y.; Mukai, M.; Ueda, J.; Muraki, M.; Stasevich, T. J.; Horikoshi, N.; Kujirai, T.; Kita, H.; Kimura, T.; Hira, S.; Okada, Y.; Hayashi-Takanaka, Y.; Obuse, C.; Kurumizaka, H.; Kawahara, A.; Yamagata, K.; Nozaki, N.; Kimura, H., Genetically encoded system to track histone modification in vivo. *Sci Rep* **2013**, *3*, 2436.
39. Tanenbaum, M. E.; Gilbert, L. A.; Qi, L. S.; Weissman, J. S.; Vale, R. D., A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **2014**, *159* (3), 635-46.
40. Hu, S.; Shively, L.; Raubitschek, A.; Sherman, M.; Williams, L. E.; Wong, J. Y.; Shively, J. E.; Wu, A. M., Minibody: A novel engineered anti-carcinoembryonic antigen antibody fragment (single-chain Fv-CH3) which exhibits rapid, high-level targeting of xenografts. *Cancer Res* **1996**, *56* (13), 3055-61.
41. Shahied, L. S.; Tang, Y.; Alpaugh, R. K.; Somer, R.; Greenspon, D.; Weiner, L. M., Bispecific minibodies targeting HER2/neu and CD16 exhibit improved tumor lysis when placed in a divalent tumor antigen binding format. *J Biol Chem* **2004**, *279* (52), 53907-14.
42. Lepin, E. J.; Leyton, J. V.; Zhou, Y.; Olafsen, T.; Salazar, F. B.; McCabe, K. E.; Hahm, S.; Marks, J. D.; Reiter, R. E.; Wu, A. M., An affinity matured minibody for PET imaging of prostate stem cell antigen (PSCA)-expressing tumors. *Eur J Nucl Med Mol Imaging* **2010**, *37* (8), 1529-38.

43. Knowles, S. M.; Zettlitz, K. A.; Tavaré, R.; Rochefort, M. M.; Salazar, F. B.; Stout, D. B.; Yazaki, P. J.; Reiter, R. E.; Wu, A. M., Quantitative immunoPET of prostate cancer xenografts with ⁸⁹Zr- and ¹²⁴I-labeled anti-PSCA A11 minibody. *J Nucl Med* **2014**, *55* (3), 452-9.
44. Han, T.; Abdel-Motal, U. M.; Chang, D. K.; Sui, J.; Muvaffak, A.; Campbell, J.; Zhu, Q.; Kupper, T. S.; Marasco, W. A., Human anti-CCR4 minibody gene transfer for the treatment of cutaneous T-cell lymphoma. *PLoS One* **2012**, *7* (9), e44455.
45. Lee, Y. K.; Kim, K. S.; Kim, J. S.; Baek, J. E.; Park, S. I.; Jeong, H. Y.; Yoon, S. S.; Jung, K. C.; Song, H. G.; Park, Y. S., Leukemia-specific siRNA delivery by immunonanoplexes consisting of anti-JL1 minibody conjugated to oligo-9 Arg-peptides. *Mol Cells* **2010**, *29* (5), 457-62.
46. Abdel-Motal, U. M.; Sarkis, P. T.; Han, T.; Pudney, J.; Anderson, D. J.; Zhu, Q.; Marasco, W. A., Anti-gp120 minibody gene transfer to female genital epithelial cells protects against HIV-1 virus challenge in vitro. *PLoS One* **2011**, *6* (10), e26473.
47. Holliger, P.; Prospero, T.; Winter, G., "Diabodies": small bivalent and bispecific antibody fragments. *Proc Natl Acad Sci U S A* **1993**, *90* (14), 6444-8.
48. Perisic, O.; Webb, P. A.; Holliger, P.; Winter, G.; Williams, R. L., Crystal structure of a diabody, a bivalent antibody fragment. *Structure* **1994**, *2* (12), 1217-26.
49. Muchekehr, R.; Liu, D.; Horn, M.; Campbell, L.; Del Rosario, J.; Bacica, M.; Moskowitz, H.; Osothprarop, T.; Dirksen, A.; Doppalapudi, V.; Kaspar, A.; Pirie-Shepherd, S. R.; Coronella, J., The Effect of Molecular Weight, PK, and Valency on Tumor Biodistribution and Efficacy of Antibody-Based Drugs. *Transl Oncol* **2013**, *6* (5), 562-72.

50. Todorovska, A.; Roovers, R. C.; Dolezal, O.; Kortt, A. A.; Hoogenboom, H. R.; Hudson, P. J., Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *J Immunol Methods* **2001**, *248* (1-2), 47-66.
51. Girgis, M. D.; Kenanova, V.; Olafsen, T.; McCabe, K. E.; Wu, A. M.; Tomlinson, J. S., Anti-CA19-9 diabody as a PET imaging probe for pancreas cancer. *J Surg Res* **2011**, *170* (2), 169-78.
52. McCabe, K. E.; Liu, B.; Marks, J. D.; Tomlinson, J. S.; Wu, H.; Wu, A. M., An engineered cysteine-modified diabody for imaging activated leukocyte cell adhesion molecule (ALCAM)-positive tumors. *Mol Imaging Biol* **2012**, *14* (3), 336-47.
53. Reddy, S.; Shaller, C. C.; Doss, M.; Shchaveleva, I.; Marks, J. D.; Yu, J. Q.; Robinson, M. K., Evaluation of the anti-HER2 C6.5 diabody as a PET radiotracer to monitor HER2 status and predict response to trastuzumab treatment. *Clin Cancer Res* **2011**, *17* (6), 1509-20.
54. Adams, G. P.; Shaller, C. C.; Dadachova, E.; Simmons, H. H.; Horak, E. M.; Tesfaye, A.; Klein-Szanto, A. J.; Marks, J. D.; Brechbiel, M. W.; Weiner, L. M., A single treatment of yttrium-90-labeled CHX-A"-C6.5 diabody inhibits the growth of established human tumor xenografts in immunodeficient mice. *Cancer Res* **2004**, *64* (17), 6200-6.
55. Ebbinghaus, C.; Ronca, R.; Kaspar, M.; Grabulovski, D.; Berndt, A.; Kosmehl, H.; Zardi, L.; Neri, D., Engineered vascular-targeting antibody-interferon-gamma fusion protein for cancer therapy. *Int J Cancer* **2005**, *116* (2), 304-13.
56. Olafsen, T.; Cheung, C. W.; Yazaki, P. J.; Li, L.; Sundaresan, G.; Gambhir, S. S.; Sherman, M. A.; Williams, L. E.; Shively, J. E.; Raubitschek, A. A.; Wu, A. M., Covalent disulfide-linked anti-CEA diabody allows site-specific conjugation and radiolabeling for tumor targeting applications. *Protein Eng Des Sel* **2004**, *17* (1), 21-7.

57. Dumoulin, M.; Conrath, K.; Van Meirhaeghe, A.; Meersman, F.; Heremans, K.; Frenken, L. G.; Muyldermans, S.; Wyns, L.; Matagne, A., Single-domain antibody fragments with high conformational stability. *Protein Sci* **2002**, *11* (3), 500-15.
58. Pain, C.; Dumont, J.; Dumoulin, M., Camelid single-domain antibody fragments: Uses and prospects to investigate protein misfolding and aggregation, and to treat diseases associated with these phenomena. *Biochimie* **2015**, *111*, 82-106.
59. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013**, *82*, 775-97.
60. Helma, J.; Cardoso, M. C.; Muyldermans, S.; Leonhardt, H., Nanobodies and recombinant binders in cell biology. *J Cell Biol* **2015**, *209* (5), 633-44.
61. Sircar, A.; Sanni, K. A.; Shi, J.; Gray, J. J., Analysis and modeling of the variable region of camelid single-domain antibodies. *J Immunol* **2011**, *186* (11), 6357-67.
62. De Genst, E.; Silence, K.; Decanniere, K.; Conrath, K.; Loris, R.; Kinne, J.; Muyldermans, S.; Wyns, L., Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc Natl Acad Sci U S A* **2006**, *103* (12), 4586-91.
63. Gutierrez, C.; Schiff, R., HER2: biology, detection, and clinical implications. *Arch Pathol Lab Med* **2011**, *135* (1), 55-62.
64. Yan, M.; Schwaederle, M.; Arguello, D.; Millis, S. Z.; Gatalica, Z.; Kurzrock, R., HER2 expression status in diverse cancers: review of results from 37,992 patients. *Cancer Metastasis Rev* **2015**, *34* (1), 157-64.
65. Vaneycken, I.; Devoogdt, N.; Van Gassen, N.; Vincke, C.; Xavier, C.; Wernery, U.; Muyldermans, S.; Lahoutte, T.; Caveliers, V., Preclinical screening of anti-HER2 nanobodies for molecular imaging of breast cancer. *FASEB J* **2011**, *25* (7), 2433-46.

66. Kijanka, M. M.; van Brussel, A. S.; van der Wall, E.; Mali, W. P.; van Diest, P. J.; van Bergen En Henegouwen, P. M.; Oliveira, S., Optical imaging of pre-invasive breast cancer with a combination of VHHs targeting CAIX and HER2 increases contrast and facilitates tumour characterization. *EJNMMI Res* **2016**, *6* (1), 14.
67. Gray, M. A.; Tao, R. N.; DePorter, S. M.; Spiegel, D. A.; McNaughton, B. R., A Nanobody Activation Immunotherapeutic that Selectively Destroys HER2-Positive Breast Cancer Cells. *Chembiochem* **2016**, *17* (2), 155-8.
68. Pruszyński, M.; Koumariānou, E.; Vaidyanathan, G.; Revets, H.; Devoogdt, N.; Lahoutte, T.; Zalutsky, M. R., Targeting breast carcinoma with radioiodinated anti-HER2 Nanobody. *Nucl Med Biol* **2013**, *40* (1), 52-9.
69. Farah, F. S., Natural antibodies specific to the 2,4-dinitrophenyl group. *Immunology* **1973**, *25* (2), 217-26.
70. Parker, C. G.; Domaoal, R. A.; Anderson, K. S.; Spiegel, D. A., An antibody-recruiting small molecule that targets HIV gp120. *J Am Chem Soc* **2009**, *131* (45), 16392-4.
71. McEnaney, P. J.; Parker, C. G.; Zhang, A. X.; Spiegel, D. A., Antibody-recruiting molecules: an emerging paradigm for engaging immune function in treating human disease. *ACS Chem Biol* **2012**, *7* (7), 1139-51.
72. Rothbauer, U.; Zolghadr, K.; Tillib, S.; Nowak, D.; Schermelleh, L.; Gahl, A.; Backmann, N.; Conrath, K.; Muyldermans, S.; Cardoso, M. C.; Leonhardt, H., Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods* **2006**, *3* (11), 887-9.
73. Kubala, M. H.; Kovtun, O.; Alexandrov, K.; Collins, B. M., Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein Sci* **2010**, *19* (12), 2389-401.

74. Kirchhofer, A.; Helma, J.; Schmidthals, K.; Frauer, C.; Cui, S.; Karcher, A.; Pellis, M.; Muyldermans, S.; Casas-Delucchi, C. S.; Cardoso, M. C.; Leonhardt, H.; Hopfner, K. P.; Rothbauer, U., Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* **2010**, *17* (1), 133-8.
75. Bruce, V. J.; Lopez-Islas, M.; McNaughton, B. R., Resurfaced cell-penetrating nanobodies: A potentially general scaffold for intracellularly targeted protein discovery. *Protein Sci* **2016**, *25* (6), 1129-37.
76. Braun, M. B.; Traenkle, B.; Koch, P. A.; Emele, F.; Weiss, F.; Poetz, O.; Stehle, T.; Rothbauer, U., Peptides in headlock - a novel high-affinity and versatile peptide-binding nanobody for proteomics and microscopy. *Sci Rep-Uk* **2016**, *6*.
77. Bruce, V. J.; McNaughton, B. R., Evaluation of Nanobody Conjugates and Protein Fusions as Bioanalytical Reagents. *Anal Chem* **2017**, *89* (7), 3819-3823.
78. Koide, A.; Bailey, C. W.; Huang, X.; Koide, S., The fibronectin type III domain as a scaffold for novel binding proteins. *J Mol Biol* **1998**, *284* (4), 1141-51.
79. Stockbridge, R. B.; Koide, A.; Miller, C.; Koide, S., Proof of dual-topology architecture of Fluc F- channels with monobody blockers. *Nat Commun* **2014**, *5*, 5120.
80. Turman, D. L.; Nathanson, J. T.; Stockbridge, R. B.; Street, T. O.; Miller, C., Two-sided block of a dual-topology F- channel. *Proc Natl Acad Sci U S A* **2015**, *112* (18), 5697-701.
81. Stockbridge, R. B.; Kolmakova-Partensky, L.; Shane, T.; Koide, A.; Koide, S.; Miller, C.; Newstead, S., Crystal structures of a double-barrelled fluoride ion channel. *Nature* **2015**, *525* (7570), 548-51.

82. Guntas, G.; Lewis, S. M.; Mulvaney, K. M.; Cloer, E. W.; Tripathy, A.; Lane, T. R.; Major, M. B.; Kuhlman, B., Engineering a genetically encoded competitive inhibitor of the KEAP1-NRF2 interaction via structure-based design and phage display. *Protein Eng Des Sel* **2016**, *29* (1), 1-9.
83. Park, S. H.; Park, S.; Kim, D. Y.; Pyo, A.; Kimura, R. H.; Sathirachinda, A.; Choy, H. E.; Min, J. J.; Gambhir, S. S.; Hong, Y., Isolation and Characterization of a Monobody with a Fibronectin Domain III Scaffold That Specifically Binds EphA2. *PLoS One* **2015**, *10* (7), e0132976.
84. Tanaka, S.; Takahashi, T.; Koide, A.; Ishihara, S.; Koikeda, S.; Koide, S., Monobody-mediated alteration of enzyme specificity. *Nat Chem Biol* **2015**, *11* (10), 762-4.
85. Esvelt, K. M.; Carlson, J. C.; Liu, D. R., A system for the continuous directed evolution of biomolecules. *Nature* **2011**, *472* (7344), 499-503.
86. Badran, A. H.; Guzov, V. M.; Huai, Q.; Kemp, M. M.; Vishwanath, P.; Kain, W.; Nance, A. M.; Evdokimov, A.; Moshiri, F.; Turner, K. H.; Wang, P.; Malvar, T.; Liu, D. R., Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance. *Nature* **2016**, *533* (7601), 58-63.
87. Sassoon, I.; Blanc, V., Antibody-drug conjugate (ADC) clinical pipeline: a review. *Methods Mol Biol* **2013**, *1045*, 1-27.

CHAPTER FIVE

Evaluation of Nanobody Conjugates and Protein Fusions as Bioanalytical Reagents⁵

5.1 Introduction

Enzyme-linked immunosorbent assay (ELISA), flow cytometry, and Western blot are common bioanalytical techniques. Successful execution traditionally requires the use of one or more commercially available antibody – small-molecule dyes or antibody – reporter protein conjugates that recognize relatively short peptide tags (<15 amino acids). However, the size of antibodies and their molecular complexity (by virtue of post-translational disulfide formation and glycosylation) typically require either expression in mammalian cells or purification from immunized mammals. As discussed in Chapter Four, the preparation and purification of chemical dye – or reporter protein – antibody conjugates is often complicated and expensive and not common place in academic laboratories. In response, researchers have developed comparatively simpler protein scaffolds for macromolecular recognition, which can be expressed with relative ease in *E. coli* and can be evolved to bind virtually any target. Examples of these reagents are thoroughly discussed in Chapter Four. One such illustration includes nanobodies, a minimalist scaffold generated from camelid-derived heavy-chain IgGs. A multitude of nanobodies have been evolved to recognize a diverse array of targets, including a short peptide. Here, this peptide tag (termed BC2T) and BC2 nanobody-dye conjugates or reporter protein fusions are evaluated in ELISA, flow cytometry, and

⁵ Adapted from: Bruce, V.J. and McNaughton B.R., *Analytical Chemistry*, **2017**, 89, 3819.

Western blot experiments and compared to analogous experiments using commercially available antibody-conjugate/peptide tag pairs (**Figure 5.1**). Collectively, the utility and practicality of nanobody-based reagents in bioanalytical chemistry is demonstrated.

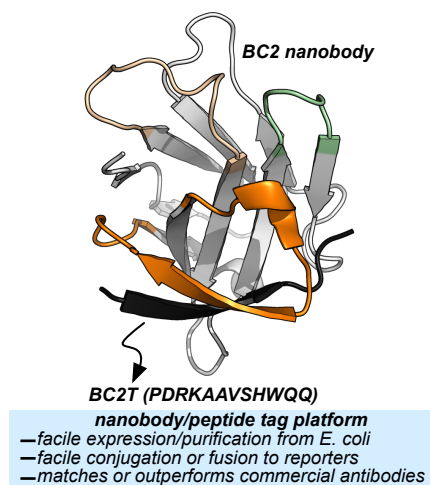


Figure 5. 1 Crystal structure of BC2 nanobody interacting with its BC2T epitope. (BC2 nanobody, PDB: 5IVN).

A multitude of bioanalytical techniques and sensor platforms rely on monoclonal antibodies with relevant examples even being utilized throughout this thesis. These antibodies are principally immunoglobulins of isotype G, IgG, and are depicted in **Figure 5.2A**.¹⁻⁴ Antibodies are large proteins (~150 kDa) that can be evolved *in vitro*, or generated by immunization, to recognize virtually any small molecule or biopolymer target. Techniques such as enzyme-linked immunosorbent assay (ELISA)⁵, flow cytometry⁴, and Western blot³ commonly rely on monoclonal antibodies that bind a small (<15 amino acid) peptide. When conjugated to a small-molecule dye or antibody—reporter protein fusion, recognition of proteins containing the peptide “tag” can occur, often in complex biological environments. Common peptide tags, for which excellent commercial antibodies and antibody – reporter conjugates exist, include FLAG⁶, myelocytomatosis viral oncogene (*myc*)⁷, synthetic streptavidin binding Strep-tag⁸ and influenza hemmagglutinin (HA)⁹.

While full-length IgG antibodies are very large (~150 kDa), target recognition is achieved within a relatively small region, termed the fragment antigen-binding (Fab) region (**Figure 5.2A**).¹⁰ Fab consists of a constant light-chain (C_L) and a variable light-chain (V_L) domain, linked to the constant ($CH1$) and the variable heavy-chain (V_H) domains. When folded properly, six solvent-exposed loops from V_L and V_H are displayed, which participate in target recognition. Collectively, these loops are referred to as the complementary determining regions (CDRs). Historically, antibodies have been a reagent of choice in bioanalytical techniques and sensor platforms (largely out of necessity); however, their size and complexity require isolation from mammalian cells or immunized mammals (principally goat, mouse, or rabbit).¹ This relatively complicated production greatly adds to the cost of antibody-based reagents, which has negative consequences in basic research and commercial diagnostics development and application.¹ Moreover, the inability of most academic laboratories to express and purify full-length antibodies, and chemically conjugate them to chemical dyes or reporter proteins, makes it challenging to prepare reagents “in house”.

Chapter Four discusses the challenges encountered with antibody-based reagents and the development of non-immuno-globulin proteins, or minimalist forms of structurally simpler immunoglobulins, as scaffolds for tailored recognition.¹⁰ Many of these scaffolds mimic the structure of IgG Fab V_H but are comparatively robust and simple to express as recombinant proteins in *E. coli*. One such scaffold is derived from heavy-chain IgGs (hcIgGs, **Figure 5.2B**). hcIgGs are produced in camelids, and in contrast to IgGs produced in other mammals, they lack a light chain. Thus, recognition is achieved through a single V_H domain, as opposed to the combination of V_H and V_L domains in IgG. In isolation, the V_H domain of hcIgG is referred to as a “nanobody”, a small (~15 kDa) protein that can be expressed in *E. coli* with relative ease, and

evolved in the laboratory to recognize a diverse array of targets, through interactions involving one or more CDR loops (CDR 1–3, **Figure 5.2C**).^{10–11} In contrast to full-length antibodies and their fragments, nanobodies often express well (>20 mg/L), as a folded and soluble recombinant protein, from *E. coli*. By virtue of their stability and ready expression, purification, and manipulation, researchers have used nanobodies in numerous applications, including therapeutic discovery¹², bioimaging¹³, and sensors^{14–15}. Rothbauer and co-workers recently reported a nanobody referred to as BC2 that binds a short peptide (PDRKAAVSHWQQ, referred to herein as BC2 tag, BC2T) with excellent affinity ($K_D \sim 1.4$ nM) and selectivity, principally through interactions involving CDR 3 (**Figure 5.2C**).¹⁶ Here, the utility of the BC2/BC2T recognition platform is evaluated in common bioanalytical techniques (ELISA, flow cytometry, and Western blot). Throughout, outcomes from the BC2 nanobody/BC2T platform are compared to those from analogous experiments using commercially available antibody–reporter conjugates and their peptide binding partners.

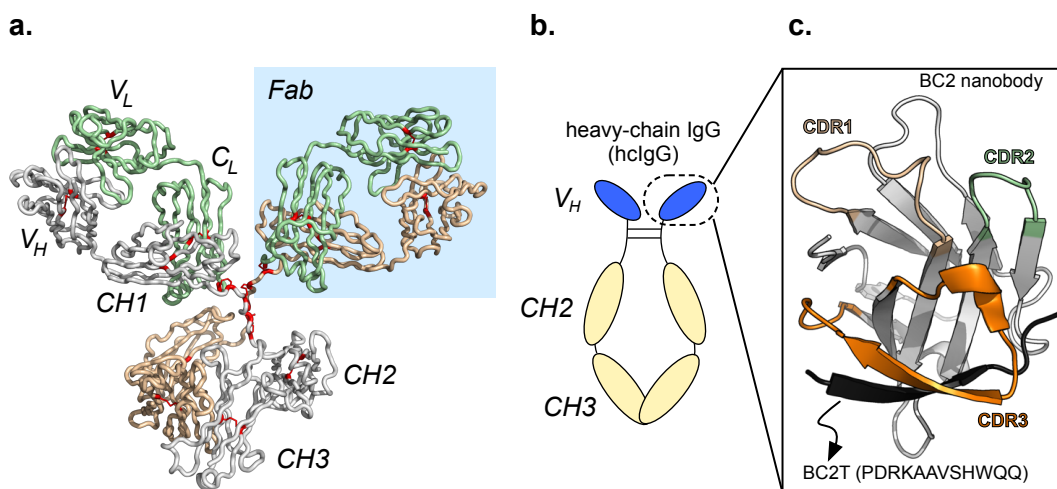


Figure 5. 2 Comparisons of Antibodies and Nanobodies. (a) Structure of IgG. Disulfide bonds are highlighted in red. Constant heavy-chain region 1 (CH1), constant light-chain (C_L), variable light-chain (V_L), variable heavy-chain (V_H), and fragment antigen-binding (Fab) regions are highlighted with a blue background (PDB: 1IGY). (b) Architecture of a heavy-chain IgG (hclgG), consisting of two heavy chains (CH3, CH2, V_H) connected by disulfide bonds in the hinge region. The “nanobody subunit is circled. (c) Structure of the recently reported nanobody BC2, bound to its peptide tag (BC2T, PDB: 5IVN).

5.2 Enzyme-Linked Immunosorbent Assay (ELISA)

ELISA typically requires (1) immobilization of a protein (“protein a”) onto a surface; (2) incubation with a binding partner (“protein b”) equipped with a small peptide tag; (3) treatment with an antibody–reporter protein conjugate, which recognizes the peptide tag and generates a signal following addition of a small-molecule substrate (**Figure 5.3A**). Horseradish peroxidase (HRP) is commonly used as a reporter protein.¹⁷ Here, a direct comparison between the BC2/BC2T platform and commercially available antibodies that bind the myc tag (EQKLISEEDL) or His6 (HHHHHH) is provided.

First, green fluorescent protein (GFP) was immobilized onto the surface of a multiwell plate. Following a washing step, GFP-coated wells were treated with buffer (NT), HRX-BC2T (which has no appreciable affinity for GFP), or a GFP-binding nanobody-His6 fusion protein (GFPnb-His6), which tightly binds GFP ($K_D \sim 1$ nM.^{13, 18} After washing steps to remove unbound material, wells were incubated with a commercially available anti-His6 antibody-HRP conjugate and HRP substrate. Unsurprisingly, no appreciable signal is observed in wells incubated with HRX-BC2T (indicating no interaction between HRX and GFP, **Figure 5.2B**, black). However, strong signal is generated in GFP immobilized wells following treatment with GFPnb-His6 and subsequent incubation with anti-His6-HRP and HRP substrate (**Figure 5.3B**, black). Similarly, when GFP immobilized wells are treated with buffer (NT), HRX-BC2T, or GFPnb-His6, no appreciable signal is observed after subsequent incubation with anti-*myc*-HRP and HRP substrate (**Figure 5.3B**, gray). However, signal that compares favorably to the analogous His6/anti-His6 experiment is observed in wells that contain immobilized GFP, following treatment with GFPnb-*myc* tag and subsequent incubation with anti-*myc*-HRP and HRP substrate (**Figure 5.2B**, gray). Satisfyingly, signal that compares favorably to the analogous experiments described above is observed when

wells containing immobilized GFP are treated with GFPnb-BC2T and subsequently incubated with “in house” prepared BC2nb-HRP conjugate and HRP substrate (**Figure 5.3B**, white, **Figure S5.1**). As expected, no appreciable signal was observed when GFP containing wells were treated with buffer (NT), HRX-BC2T, or GFPnb-His6, following subsequent treatment with BC2nb-HRP and HRP substrate (**Figure 5.3B**, white). Additionally, no appreciable signal was observed in wells lacking immobilized GFP (**Figure S5.1**).

Fundamentally, this paper aims to accentuate practical remedies afforded by the use of nanobody-based reagents in bioanalytical chemistry. It therefore went unnoticed that chemical conjugation of HRP to BC2 is likely an impediment to the broad use of this reagent, including laboratories without experience in bioconjugation. A more practical solution is expression of the BC2 nanobody as a fusion to a reporter protein. Unfortunately, the BC2 nanobody–HRP fusion does not express as a soluble protein in *E. coli*. However, a recently reported bioluminescent “nanoluciferase” protein (nLuc),¹⁹ developed by Promega, expresses as a fusion to BC2 nanobody (**Figure S5.2**). Satisfyingly, the BC2 nanobody–nLuc performed well in our ELISA analysis. First, biotinylated GFP was immobilized onto streptavidin-coated plates. Wells containing immobilized GFP were then incubated with buffer (NT), HRX-BC2T, GFPnb, or GFPnb-BC2T. Following washing steps to remove unbound material, wells were treated with the BC2- nLuc fusion protein, washed again, and then treated with the nLuc substrate (“NanoGlo”). As expected, no appreciable signal was generated in wells containing immobilized GFP but incubated with either HRX or GFP-binding nanobody lacking the BC2T peptide (**Figure 5.3C**, red and orange, respectively). In contrast, we observed robust signal in lanes containing immobilized GFP in complex with the GFP-binding nanobody genetically fused to the BC2T peptide (**Figure 5.3C**, green). When immobilized GFP was treated with a solution containing equal parts GFP-binding nanobody-BC2T

peptide and GFP-binding nanobody (without the tag), a ~50% decrease in luminescence is observed, compared to wells treated with only the GFP-binding nanobody equipped with the BC2T peptide (**Figure 5.3C**, blue). In contrast, no appreciable signal was observed in wells that were treated identically but lack immobilized GFP (**Figure S5.2**). Collectively, these data show that an “in house” prepared BC2 nanobody–nLuc fusion protein, when paired with binding partners containing the BC2 tag, is an excellent reagent for ELISA.

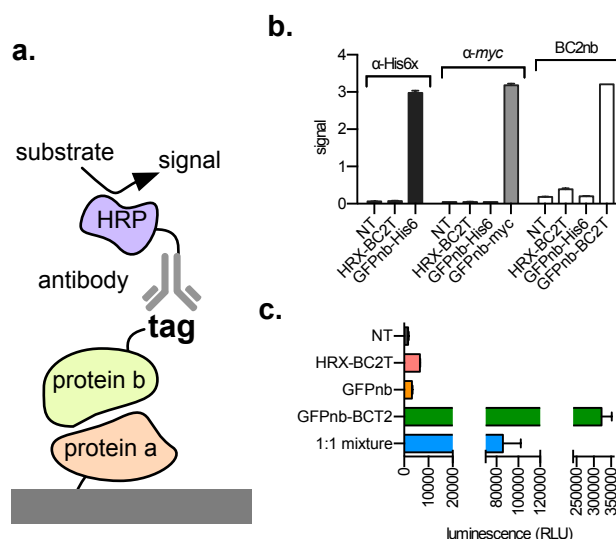


Figure 5. 3 Analysis of Nanobodies in ELISAs. (a) Scheme of an enzyme-linked immunosorbent assay (ELISA). (b) ELISA data: immobilized GFP is treated with buffer (NT) and HRX-BC2T, GFPnb-His6, GFPnb-myc, or GFPnb-BC2T, then anti-His6-HRP, anti-myc-HRP, or the BC2nb-HRP conjugate, and HRP substrate. Signal is the observed absorbance at 655 nm. (c) ELISA data: GFP was immobilized onto streptavidin coated plates and then treated with buffer (NT, black), HRX-BC2T (red), GFP nanobody (GFPnb, orange), GFPnb-BC2T (blue), followed by nLuc substrate. All experiments were performed in triplicate. Error bars represent standard deviation of three experiments. α = anti; NT = no treatment. RLU = relative luminescence units.

5.3 Flow Cytometry Applications

The BC2 nanobody/BC2T platform was next evaluated in the context of flow cytometry, a commonly used technique to evaluate protein–protein and protein–nucleic acid interactions on the surface of yeast or bacteria, and enrichment of binders from a protein library by fluorescence

activated cell sorting (FACS). In a typical flow cytometry experiment, bacteria²⁰⁻²² or yeast²³⁻²⁵ display a peptide or protein that is flanked by a peptide tag recognized by a commercial antibody–fluorescent dye conjugate. Interaction between the tag and antibody–reporter conjugate allows researchers to quantitate display efficiency. Concomitantly, the peptide or protein displaying cells are treated with a binding target that is also fluorescently tagged.

Traditionally, yeast display efficiency has been measured using a commercially available antibody–dye conjugate that binds to either an N-terminal HA tag or a C-terminal *myc* tag.²³⁻²⁵ Bacterial display efficiency on *E. coli* is typically measured using a commercial antibody that binds to a C-terminal *myc* tag. To permit direct comparative analysis, bacteria (*E. coli*) were engineered to display a small (~15 kDa) well behaved protein (monomeric streptavidin, mSA2), with flanking N-terminal and C-terminal BC2T and *myc* tags, respectively (**Figure 5.4A**). Yeast were engineered to display a HA-mSA2-BC2T-*myc* fusion (**Figure 5.4B**).

For *E. coli*, cells were induced to express the displayed protein/tag fusion (as a fusion to OmpX, an *E. coli* cell surface protein typically used for bacterial display) and then treated with a commercially available anti-*myc*-FITC antibody–fluorescent dye conjugate, an “in house” prepared BC2 nanobody-Cy5 conjugate (BC2nb-Cy5), or a BC2 nanobody-GFP fusion protein (BC2nb-GFP). Following washing steps to remove unbound material, cells were analyzed by flow cytometry, using a laser/detection channel specific to either Cy5 or FITC (GFP). Both the BC2nb-Cy5 conjugate and BC2nb-GFP fusion compared favorably to the anti-*myc*-FITC antibody–fluorescent dye conjugate (~98% display efficiency for each, **Figure 5.4C**). Co-treatment with equal parts anti-*myc*-FITC and BC2nb-Cy5 show essentially identical fluorescence (recognition of their respective displayed peptide tag (**Figure 5.4C**)). Representative flow cytometry histograms are provided in **Figure S5.3** and **Table S5.1**.

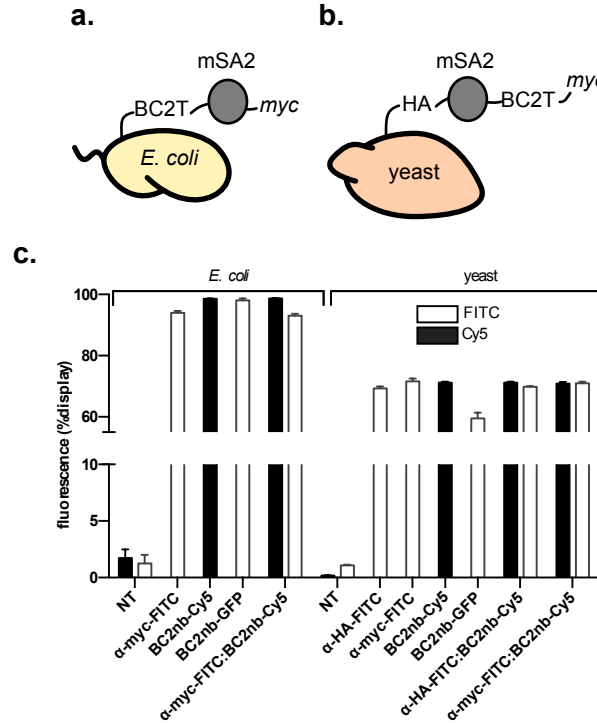


Figure 5. 4 Evaluation of Nanobodies for use in Flow Cytometry Experiments. (a) Representation of *E. coli* engineered for flow cytometry experiments. (b) Representation of yeast engineered for flow cytometry experiments. (c) Flow cytometry detection of displayed monomeric streptavidin (mSA2) on the surface of *E. coli* or yeast, as determined by commercially available antibody α -myc-FITC (for *E. coli*), nanobody reagents BC2nb-Cy5 or BC2nb-GFP (for *E. coli*), commercially available antibodies α -myc-FITC or α -HA-FITC (for yeast), or nanobody reagents BC2nb-Cy5 or BC2nb-GFP (for yeast). All experiments were performed in triplicate. Error bars represent standard deviation of three experiments. α = anti; NT = no treatment.

For yeast, cells were induced to express the displayed protein/tag fusion at the C-terminus of Aga2 (a yeast cell surface protein typically used for yeast display) and then treated with a commercially available anti-myc-FITC, anti-HA-FITC antibody–fluorescent dye conjugate, BC2nb-Cy5 conjugate, or BC2nb-GFP fusion. Again, the nanobody reagents compared favorably to commercially available antibody reagents. Individual treatment or co-treatment with equal parts anti-myc-FITC and BC2nb-Cy5 or anti-HA-FITC and BC2nb-Cy5 show essentially identical

fluorescence (recognition of their respective displayed peptide tag, **Figure 5.4C**). Representative flow cytometry histograms are provided in **Figure S5.3** and **Table S5.2**.

5.4 Western Blot Applications

As a final evaluative measure, the utility of the BC2 nanobody/BC2T platform was assessed in a Western blot, a commonly used technique to measure the presence of a specific protein (such as a tagged protein) in cell lysate. Execution of a Western blot typically requires: (1) denaturation of proteins from cell lysate; (2) separation of proteins based on their size via SDS-polyacrylamide gel electrophoresis (SDS-PAGE); (3) electrophoretic transfer of separated proteins to a membrane; (4) treatment of the protein-bound membrane with a primary antibody that recognizes either a specific protein or a specific peptide tag; (5) treatment with a secondary antibody–dye conjugate, which serves to illuminate the primary antibody-bound protein. To function in this context, the BC2 nanobody must recognize the BC2T tag following a chemical denaturation step (and subsequent denaturation of the protein to which it is attached). For this reason, many antibodies (and nanobodies) are not suitable for Western blot analysis.

For comparison to IR dye 790-labeled commercially available secondary antibody, an IR dye 800-labeled BC2 nanobody conjugate was prepared by reaction between a C-terminal cysteine and commercially available dye maleimide. First, 5 μ M GFP lacking the BC2T peptide, or GFP-BC2T, was run on a polyacrylamide gel, transferred to PVDF membrane, and treated with BC2nb-IR800 reagent. Only GFP-BC2T was detected, but not GFP lacking BC2T peptide, indicating that recognition relies entirely on the nanobody–tag recognition, in this context (**Figure S5.4**). Next, purified GFP-HA, GFP-*myc*, or GFP-BC2T was ran in duplicate on a polyacrylamide gel at 20, 10, 5, and 1 μ M concentrations. Following PAGE, one gel was stained by Coomassie to determine

protein purity. Proteins embedded in the other gel were transferred onto a PVDF membrane. Membranes containing GFP-HA or GFP-*myc* were first treated with commercially available anti-HA or anti-*myc* primary antibodies suggested for Western blot experiments. Next, these membranes were treated with a secondary antibody–Alexa Fluor 790 dye. Following washing steps, membranes were imaged on a Li-Cor Odyssey instrument. All three proteins (GFP-HA, GFP-*myc*, or GFP-BC2T) were found to be pure, as determined by Coomassie staining (**Figure 5.5A-C**, left gels). As expected, both anti-HA and anti-*myc* antibodies recognize HA or *myc* tagged proteins in the Western blot (**Figure 5.5 A-B**, right gels). Satisfyingly, the BC2nanobody-IR800 dye conjugate recognized GFP-BC2T with excellent potency and selectivity (**Figure 5.5C**, right gels). In fact, the BC2nb/BC2T pair generated a more robust and cleaner signal, in comparison to the HA and *myc* platforms. SDS-PAGE analysis of additional purified proteins used in this work can be found in **Figure S5.5**.

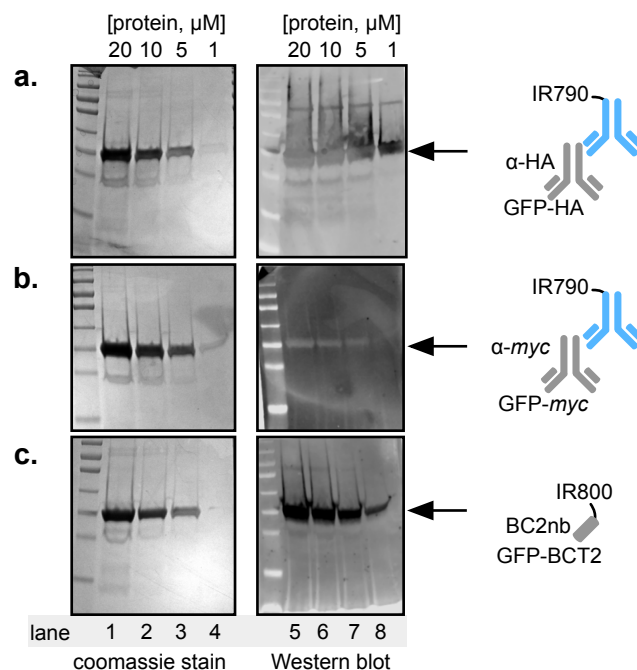


Figure 5. 5 Evaluation of Nanobodies for Western Blot Applications. (a-c, left gel) Coomassie stained polyacrylamide gels following loading with 20, 10, 5, or 1 μM GFP-HA, GFP-*myc*, or GFP-BC2T and electrophoresis. (a-c, right gel) Western blot data for GFP-HA/anti-HA, GFP-*myc*/anti-*myc*, or GFP-BC2T/BC2nb pairs respectively. α = anti.

5.5 Conclusion

In conclusion, antibodies and their conjugates play a central role in a multitude of bioanalytical methods and sensor platforms. However, their cost and complexity add to challenges with their use. Researchers have developed minimalist protein architectures, which mimic structural features found in the antigen binding region of antibodies. One minimalist scaffold is the nanobody, a camelid-derived protein that can be evolved to bind virtually any target, including relatively short peptides. In contrast to antibodies, nanobodies express well in *E. coli* and can be easily manipulated, such as conjugation to a small-molecule dye or genetic fusion to a reporter protein. Collectively, these features make nanobody-based reagents an attractive alternative to antibodies and their conjugates. Using a recently reported BC2 nanobody/BC2T peptide tag pair and “in house” prepared nanobody conjugates and fusion proteins, comparative analysis to commercially available antibodies and antibody conjugates has been conducted. In every platform tested (ELISA, flow cytometry, and Western blot), nanobody-based reagents compare favorably to, or outperform, antibody-based reagents. We hope these findings encourage the use of nanobody-based reagents in bioanalytical methods and lead to the evolution of new nanobody/peptide tag binding pairs.

REFERENCES

1. Byrne, B.; Stack, E.; Gilmartin, N.; O'Kennedy, R., Antibody-based sensors: principles, problems and potential for detection of pathogens and associated toxins. *Sensors (Basel)* **2009**, *9* (6), 4407-45.
2. Holford, T. R.; Davis, F.; Higson, S. P., Recent trends in antibody based sensors. *Biosens Bioelectron* **2012**, *34* (1), 12-24.
3. Liu, Z. Q.; Mahmood, T.; Yang, P. C., Western blot: technique, theory and trouble shooting. *N Am J Med Sci* **2014**, *6* (3), 160.
4. Laerum, O. D.; Farsund, T., Clinical application of flow cytometry: a review. *Cytometry* **1981**, *2* (1), 1-13.
5. Carlsson, H. E.; Hurvell, B.; Lindberg, A. A., Enzyme-linked immunosorbent assay (ELISA) for titration of antibodies against *Brucella abortus* and *Yersinia enterocolitica*. *Acta Pathol Microbiol Scand C* **1976**, *84* (3), 168-76.
6. Einhauer, A.; Jungbauer, A., The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. *J Biochem Biophys Methods* **2001**, *49* (1-3), 455-65.
7. Evan, G. I.; Lewis, G. K.; Ramsay, G.; Bishop, J. M., Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol Cell Biol* **1985**, *5* (12), 3610-6.
8. Schmidt, T.; Skerra, A., The Strep-tag system for one-step affinity purification of proteins from mammalian cell culture. *Methods Mol Biol* **2015**, *1286*, 83-95.
9. Wilson, I. A.; Niman, H. L.; Houghten, R. A.; Cherenson, A. R.; Connolly, M. L.; Lerner, R. A., The structure of an antigenic determinant in a protein. *Cell* **1984**, *37* (3), 767-78.

10. Bruce, V. J.; Ta, A. N.; McNaughton, B. R., Minimalist Antibodies and Mimetics: An Update and Recent Applications. *Chembiochem* **2016**, *17* (20), 1892-1899.
11. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013**, *82*, 775-97.
12. Gray, M. A.; Tao, R. N.; DePorter, S. M.; Spiegel, D. A.; McNaughton, B. R., A Nanobody Activation Immunotherapeutic that Selectively Destroys HER2-Positive Breast Cancer Cells. *Chembiochem* **2016**, *17* (2), 155-8.
13. Schoonooghe, S.; Laoui, D.; Van Ginderachter, J. A.; Devoogdt, N.; Lahoutte, T.; De Baetselier, P.; Raes, G., Novel applications of nanobodies for in vivo bio-imaging of inflamed tissues in inflammatory diseases and cancer. *Immunobiology* **2012**, *217* (12), 1266-72.
14. Buchfellner, A.; Yurlova, L.; Nuske, S.; Scholz, A. M.; Bogner, J.; Ruf, B.; Zolghadr, K.; Drexler, S. E.; Drexler, G. A.; Girst, S.; Greubel, C.; Reindl, J.; Siebenwirth, C.; Romer, T.; Friedl, A. A.; Rothbauer, U., A New Nanobody-Based Biosensor to Study Endogenous PARP1 In Vitro and in Live Human Cells. *PLoS One* **2016**, *11* (3), e0151041.
15. Chen, J.; He, Q. H.; Xu, Y.; Fu, J. H.; Li, Y. P.; Tu, Z.; Wang, D.; Shu, M.; Qiu, Y. L.; Yang, H. W.; Liu, Y. Y., Nanobody medicated immunoassay for ultrasensitive detection of cancer biomarker alpha-fetoprotein. *Talanta* **2016**, *147*, 523-30.
16. Braun, M. B.; Traenkle, B.; Koch, P. A.; Emele, F.; Weiss, F.; Poetz, O.; Stehle, T.; Rothbauer, U., Peptides in headlock--a novel high-affinity and versatile peptide-binding nanobody for proteomics and microscopy. *Sci Rep* **2016**, *6*, 19211.
17. Beyzavi, K.; Hampton, S.; Kwasowski, P.; Fickling, S.; Marks, V.; Clift, R., Comparison of Horseradish-Peroxidase and Alkaline Phosphatase-Labeled Antibodies in Enzyme Immunoassays. *Ann Clin Biochem* **1987**, *24*, 145-152.

18. Kubala, M. H.; Kovtun, O.; Alexandrov, K.; Collins, B. M., Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. *Protein Sci* **2010**, *19* (12), 2389-401.
19. Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V., Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7* (11), 1848-1857.
20. Lofblom, J., Bacterial display in combinatorial protein engineering. *Biotechnol J* **2011**, *6* (9), 1115-29.
21. Daugherty, P. S., Protein engineering with bacterial display. *Curr Opin Struct Biol* **2007**, *17* (4), 474-80.
22. Laplagne, D. A.; Zylberman, V.; Ainciart, N.; Steward, M. W.; Sciutto, E.; Fossati, C. A.; Goldbaum, F. A., Engineering of a polymeric bacterial protein as a scaffold for the multiple display of peptides. *Proteins* **2004**, *57* (4), 820-8.
23. Boder, E. T.; Wittrup, K. D., Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol* **2000**, *328*, 430-44.
24. Kieke, M. C.; Shusta, E. V.; Boder, E. T.; Teyton, L.; Wittrup, K. D.; Kranz, D. M., Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc Natl Acad Sci U S A* **1999**, *96* (10), 5651-6.
25. Boder, E. T.; Wittrup, K. D., Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology* **1997**, *15* (6), 553-557.

CHAPTER SIX

Part I: Intracellular NanoBit Assembly: A Facile Method to Measure Cytosolic Residence of Delivered Proteins and Part II: Innovative Platforms for *In Vivo* Molecular Tagging and Imaging

6.1 Introduction

A major subset of my research in the McNaughton Laboratory has focused on the optimization and identification of cell penetrating peptides or whole proteins for the delivery of exogenous proteins to mammalian cells, and the evaluation of nanobodies as sensible antibody alternatives for bioanalytical assays. Concomitant with this, I have also had great interest in developing practical remedies, often resulting in straightforward and simplified methods, for assays utilized (or could have been utilized) in the first portion of this thesis. That additional work has resulted in continuing projects and manuscripts in preparation, two of which are described in this chapter.

My previously discussed projects concerning exogenous protein delivery to mammalian cells (Chapters Two and Three) report initial determination of protein delivery through the fusion of that reagent (either CPP or nanobody) with GFP and then flow cytometry and fluorescent microscopy analysis of those cells using internalized GFP as an indicator for cellular uptake. While these methods (and other methods similar to it) ultimately yield useful information, they cannot easily comment on the subcellular location of these protein deliverables, a paramount distinction if these reagents are to act on therapeutically-relevant intracellular targets located in the cytosol. In response to this challenge, Jennifer Bjerke, a current fourth-year graduate student in the

McNaughton Laboratory, has been working with me to devise a technology platform based on intracellular reassembly of a recently reported split-nanoluciferase variant (NanoBit) for measurement of cytosolic residence of delivered proteins to mammalian cells. Effort towards this technology are described in Part I of this chapter.

Part II of this chapter describes the start of collaboration work between our group and the Stasevich Laboratory of the Biochemistry Department at CSU. Work described below was completed with the assistance of Stasevich group members Tatsuya Morisaki (research scientist) and Ning Zhao (post-doctoral scientist). The Stasevich Laboratory utilizes high copy tandem repeats or amplification tags (12X tags) of short epitopes (such as FLAG-tag, DYKDDDDK) to flank proteins and sections of nucleic acids of interest for live-cell imaging. These constructs are then translated inside mammalian cells where exogenous reagents can tag the short epitopes allowing for the visualization and quantification of translation dynamics in living cells. Specifically, we are investigating expertise in the McNaughton Laboratory for the generation and characterization of biomolecular assemblies to produce ideal, orthogonal, and minimalistic reagents for investigation of *in vivo* assemblies in living cells. We believe these reagents can contribute significantly to single molecule spectroscopy. Initial work towards this goal are described in Part II of this chapter.

6.2 Part I: Intracellular NanoBit Assembly: A Facile Method to Measure Cytosolic Residence of Delivered Proteins

Proteins are increasingly used as basic research tools and therapeutics. However, as expressed in Chapter One, the inability of most proteins to cross the lipid bilayer of mammalian cells dramatically restricts their use. Unsurprisingly, methods for intracellular delivery of exogenous

proteins have received significant attention over the last decade. Multiple methods, as discussed in the first portion of this thesis, for intracellular protein delivery now exist, however, many of these routes proceed via endocytosis. Importantly, in order to engage their intracellular target, many exogenously delivered proteins must ultimately obtain residence in the cytosol. Therefore, simple and medium- to high-throughput methods to measure cytosolic residence, which rely on easy to access and relatively inexpensive reagents and/or equipment, are needed to evaluate functional delivery of exogenous proteins. Here, we report such a method, which relies on intracellular reassembly of a recently reported split-nanoluciferase variant (NanoBit) (**Figure 6.1**). The luminescent signal in this assay is only generated when a Small-Bit (SmBit) peptide tag on the delivered protein reassembles with a constitutively expressed Large-Bit (LgBit) protein. This assay can be performed at medium- to high-throughput efficiency using commercially available reagents and plate readers. We anticipate this method will prove useful to researchers interested in developing protein delivery methods, and measuring cytosolic residence of those proteins in mammalian cells.

Because proteins typically cannot appreciably penetrate mammalian cells, exogenously administered proteins are typically limited to targeted cell-surface or circulating receptors, greatly

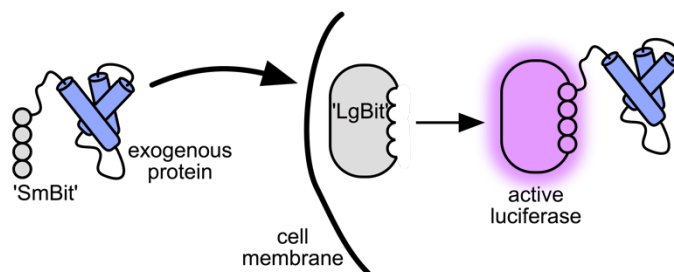


Figure 6. 1 Overview of Intracellular NanoBit Reassembly. A luminescence response (indicating cytosolic delivery) is achieved only when proteins of interest fused with the ‘SmBit’ peptide reassemble with the cytosolically expressed ‘LgBit’ subunit of NanoBit.

limiting their broader use. The challenge to achieve cell-selective, safe, and efficient protein delivery to mammalian cells has encouraged the development of a vast array of methods; and, of current, has received enormous effort in these regards. Exogenous protein reagents must be capable of reaching the cytosol in order to behave as functional therapeutics and properly engage their target. Functional analysis of intracellular protein delivery technologies must include assays that report on cytosolic residence. At current, analysis of these delivery methods lack robust and meaningful interrogation procedures or techniques that can make judgement on the interior location of these delivered macromolecules.

Current methods often rely on Western blot, flow cytometry, and fluorescence microscopy, which can be relatively low-throughput, require expensive equipment and expertise, and results in findings that are often difficult to comment on specific intracellular location. Alternatively, functional stipulations can be implemented, including assays where the protein deliverable has enzymatic activity that results in a phenotypic change. One example is arginine-grafted RNaseA cytotoxin delivery which induces apoptosis.¹ However, these methods require, (1) connection of the protein deliverable with a readable phenotypic change and (2) a great-deal of prior functional knowledge concerning that protein and its capabilities.

More general assays along these lines include the use of reporter systems such as Cre-recombinase.²⁻³ This system can comment on subcellular location because the enzyme Cre recombinase must translocate to the nucleus and recombine the reporter gene cassette to trigger expression of a fluorescent protein reporter. While this method essentially associates your protein of interest (deliverable) with phenotypic functionality, it requires the fusion of a rather large protein (Cre, ~ 38 kDa) to your protein in question as well as a cleavable linker between the two, as it has been noted that Cre fusions have little to no recombinase activity until the additional

protein moiety has been cleaved. Thus, this method results in greatly increased complexity to the system and an indirect response. These examples strongly illustrate the challenges in determining exogenous protein cellular delivery.

Because fluorescent proteins, such as GFP, can act as stable, genetically encoded fluorophores that furthermore can function in mammalian cells, assays that utilize GFP have been heavily investigated. Specifically, the Schepartz Laboratory employs GFP induction or translocation as a readout to quantify cytosolic localization of proteins.⁴⁻⁵ However this method is highly dependent on expensive instrumentation and expertise and / or does not yield itself to high-throughput analysis. Another method based on GFP fluorescence was recently reported by the Dowdy Laboratory. They employ a split-GFP peptide complementation assay to comment on the subcellular location of proteins fused to the cell penetrating peptide Tat⁶⁻⁷ and a series of synthetic endosomal escape domains (EEDs).⁸ While these systems experiences several benefits to the aforementioned methods, they still require chemical formation of the fluorescent chromophore to reveal cellular uptake, do not have an ‘on’ switch, and are fairly large in size (GFP is about a 1/3 of the size larger than NanoBit). Additionally, the size and spectral properties of GFP impose certain limitations for its use and circumstances or the desire of a readout to be something other than a fluorescent protein can be envisioned. The review of such assays highlights both the power and current limitations for determining cytosolic residence of exogenously delivered proteins. An ideal platform would include (1) minimal tags requirements, (2) a highly stable reporter with a well characterized readout, and (3) the capability of high-throughput analysis that is not reliant on expensive reagents and/or equipment.

Part I of this chapter reports the design and development of a novel assay for measuring cytosolic residence of intracellularly delivered proteins. This assay relies on intracellular

reassembly of a recently reported split-nanoluciferase variant (NanoBit).⁹ The luminescent signal in this assay is only generated when SmBit, an 11-mer amino acid tag, on the delivered protein reassembles with a constitutively expressed, cytosolic LgBit (~ 18kDa) protein. The LgBit segment has been robustly designed for stability and use as a reporter protein under relevant physiological conditions inside living cells. Additionally, this assay is highly amenable to medium- to high-throughput analysis, and relies on relatively inexpensive and commercially available reagents, and commercially available luminescence plate readers.

6.3 Development of Intracellular NanoBit Assembly

Initial development of the NanoBit technology by Promega as a complementation reporter for the measurement of protein-protein interactions in cells resulted in multiple LgBit and SmBit fragments with different binding affinities and stabilities. Because we were not interested in protein-protein interactions and desire quick intracellular reassembly if both fragments are present, we chose a combination of NanoBit complementarity fragments with the highest affinity for one another ($K_D = 0.7 \times 10^{-9}$ M). We hypothesized that a luminescence signal could be generated through exogenous delivery of the SmBit peptide when attached to a protein of interest once brought into contact with the constitutively expressed, cytosolic LgBit protein (**Figure 6.2**). To facilitate this, we cloned the LgBit portion of NanoBit into a mammalian expression vector that would express LgBit as a fusion protein with a fluorescent protein (GFP). The purpose of a fluorescent protein here is to allow comment on the transfection efficiency achieved for LgBit production. Next, we made fusion clones of the 11-mer SmBit peptide to a variety of proteins, including proteins known to penetrate mammalian cells as well as some shown not to possess those capabilities (**Table 6.1**). Thereby, if a protein is capable of reaching cytosolic residence, the SmBit

peptide would likely engage with its LgBit binding partner, thus generating measurable and quantifiable luminescence signal that can be analyzed on a plate reader.

Specifically, this system is ideal for intracellular protein delivery interrogation and benefits from a small size (total combined components ~19 kDa) and thus lower probability of steric hindrance or peptide impacting internalization, with a fully characterized reporter protein element that produces a bright luminescence readout.

6.4 Design of Live-Cell Quantitative NanoBit Luminescence Transduction Assay

Recently, we described a polycationic resurfacing strategy for the exogenous protein delivery of nanobodies to mammalian cells.¹⁰⁻¹¹ In this account, we validated cellular uptake of GFP fusions of these reagents via fluorescence microscopy and flow cytometry. Interestingly, the results from these internalized reagents did not appear as punctate foci when visualized under a fluorescent microscope, as seen with arginine grafted GFP and supercharged GFP¹²⁻¹³, suggesting these polycationic resurfaced nanobodies achieved some degree of residence in the cytosol. This result was further analyzed using Western blot analysis of digitonin extracted cell lysate – which breaks the cell surface lipid bilayer, but not endosomes. While this assay commented on the cellular

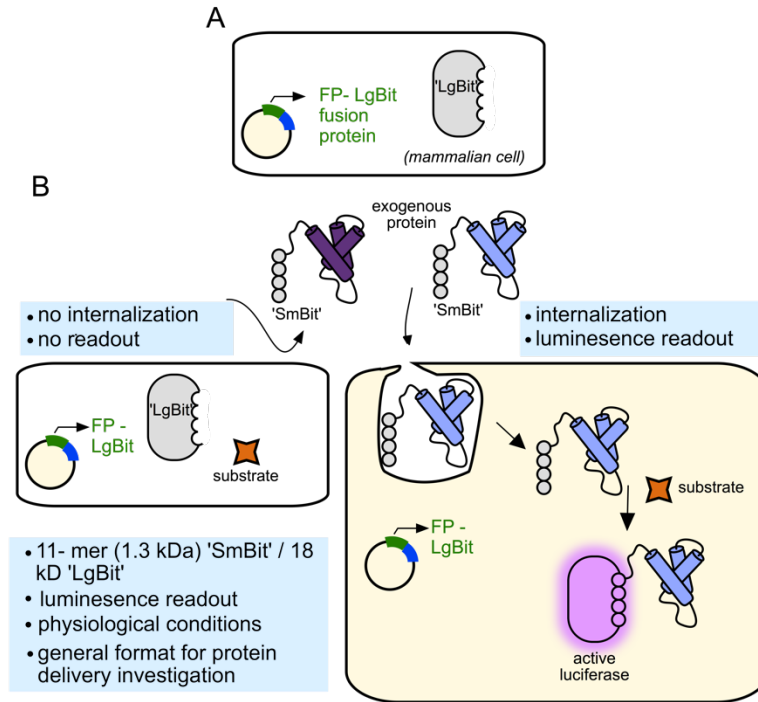


Figure 6.2 Overview of *In Vivo* Intracellular NanoBit Assay. A) Cells are transfected with a plasmid to express FP-LgBit; B) Cells are then incubated with exogenous proteins fused with SmBit. If cytosolic delivery occurs then reassembly of NanoBit will produce an active luciferase construct that can react with the cell-permeable NanoGlo small-molecule substrate to produce a luminescent signal.

location of our polycationic resurfaced nanobodies, the method cannot be done in a high-throughput format, and is often at a disadvantage from a lack robust antibodies for compartment and protein of interest identification. Specifically, we wished to compare the results gained from that analysis to our proposed NanoBit assay that can be completed at medium- to high- throughput.

First we purified proteins containing the SmBit peptide (β -lactamase nanobody and its polycationic resurfaced variant)¹⁰ and investigated if they could associate with the purified fluorescent protein (FP)-LgBit fusion protein (**Figure S6.1**). *In vitro* mixing of the purified components, outside the complex environment of mammalian cells, resulted in a robust signal of luminescence. Additionally, similar results were observed when purified proteins were mixed in a solution containing mammalian cell lysate. Mixing of nanobodies containing the Sm-Bit fusion at

various concentrations (50 – 0.05 nM) in mammalian cell lysate with an excess of FP-LgBit (1 μ M) resulted in a decrease in luminescence signal as the predetermined concentration of the Sm-Bit component decreased (**Figure 6.3**). Importantly, this suggests the luminescence signal is dependent on the concentration of the protein of interest-SmBit fusion (or how much of the exogenous protein reaches the cytoplasm). This key point endows the platform with a degree of quantification instead of a simple binary (internalization: yes or no) response. In particular, this is crucial, as enzymatic proteins such as Arg-RNase A might require much lower levels of delivery than non-enzymatic reagents.

Table 6. 1 Proteins to be assayed for delivery and cytosolic residence. All will be made as fusion proteins with SmBit-86 on the C-term.

	protein
1	β -lac nanobody
2	β -lac pcNanobody
3	GFP
4	Arg grafted GFP
5	15 GFP
6	34 GFP

Next, we administered the SmBit tagged nanobodies (β -lac wild-type and polycationic resurfaced) to Hek293 cells that had been transfected with a plasmid encoding FP-LgBit to examine if we can quantitatively determine the cytoplasmic uptake of these proteins in live cells via luminescence signal. First, we plated Hek293 cells in a 24-well plate with black sides and a clear bottom. Using standard mammalian cell transfection techniques, these cells were then made to constitutively express the cytosolic FP-LgBit protein. After a 12-hour incubation with DNA and transfection reagents, Hek293 cells were washed and an OptiMem medium solution containing

500 nM exogenous SmBit tagged protein was added to those cells for another 24-hour incubation. These cells were then washed with phosphate buffered saline solution containing 20 U/mL heparin sulfate – which has been previously shown to remove cell surface bound protein especially supercharged proteins.¹³⁻¹⁸ Following washing, these cells were then imaged on a fluorescent microscope to comment on transfection efficiency. Significant levels of GFP were observed suggesting transformation efficiency was >90% (**Figure S6.2**). Finally, a 1:50 dilution of nano-luciferase substrate (NanoGlo) in PBS was added to cells while gently shaking for ~10 min before reading luminescence.

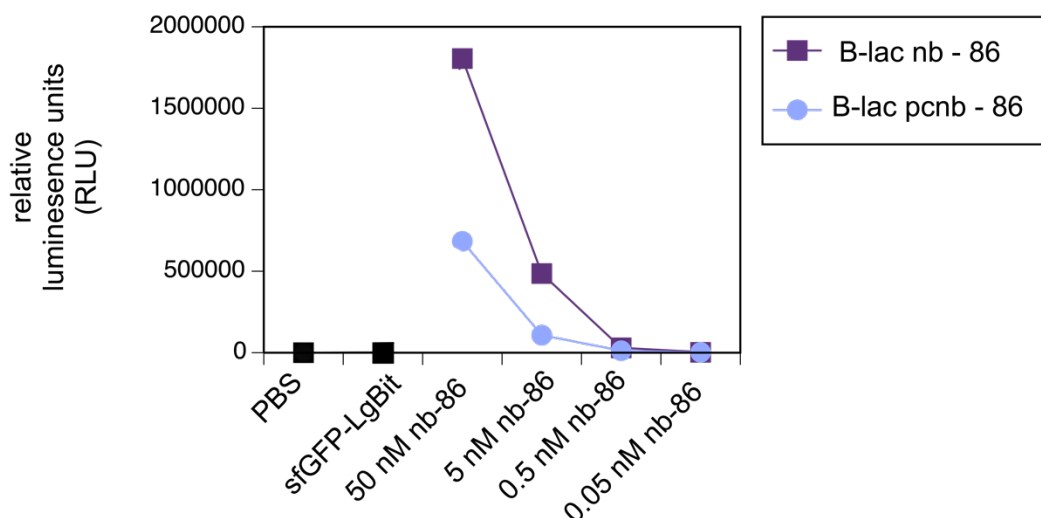


Figure 6. 3 *In Vitro* Mixing of Purified Proteins Produces Luminescence. Purified nanobody proteins fused to SmBit when reacted with sfGFP-LgBit produce luminescence upon adding substrate. As the concentration of nanobody-SmBit proteins decreases there is also a decrease in observed luminescence.

As expected, wells containing the wild-type β -lac nanobody-SmBit did not produce appreciable luminescence signal (**Figure 6.4A**) suggesting no internalization. In contrast, cells incubated with the polycationic resurfaced β -lac nanobody-SmBit produce a robust signal suggesting further validation of cytosolic access for these engineered proteins. When any individual component is missing, (non-transfected cells, non-transfected cells treated incubated

with protein, transfected cells with no protein of interest added) no observable luminescence signal is achieved (**Figure 6.4B**). Overall, this data supports our hypothesis, that intracellular NanoBit assembly can be used as a facile method to measure cytosolic residence of delivered proteins.

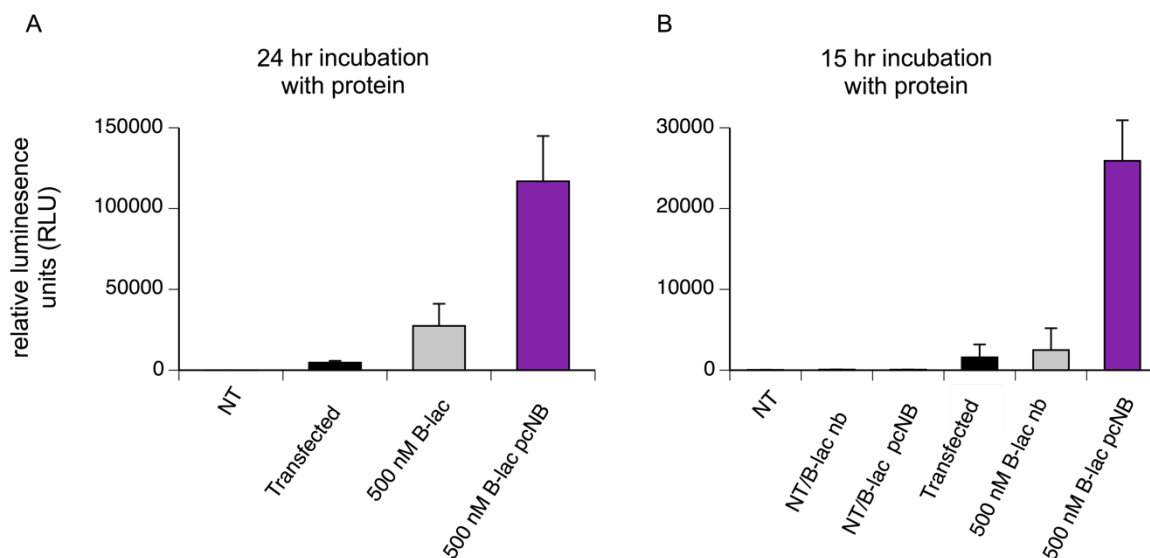


Figure 6. 4 *In Vivo* Intracellular NanoBit Assay. A) Hek293 cells incubated with exogenous proteins for 24 hours before reading; B) Hek293 cells incubated with exogenous protein for 15 hours before reading for luminescence. NT = No transfection.

Interestingly, when the same experiment was completed, but exogenous protein was only allowed ~15-hours of incubation instead of 24-hours with LgBit expressing mammalian cells, there was a notable decrease in luminescence signal from wells incubated with the polycationic resurfaced nanobody (**Figure 6.4B**). This supports an endocytic mode of uptake for these reagents and that perhaps given adequate time they are able to achieve endosomal escape allowing cytosolic localization.

6.5 Continuing Work – NanoBit

To assess the contribution of endocytic entry, we produced fusion proteins of all protein candidates listed in **Table 6.1** with an N-terminal endosomolytic peptide, aurein 1.2, recently developed by the Liu Laboratory.¹⁹ This particular sequence was initially identified because it belonged to a class of membrane-active peptides that also function as antimicrobial agents based on their ability to penetrate microbial membranes. Specifically, this 13-mer peptide (GLFDIIKKIAESF) has been shown to enhance cytosolic delivery of supercharged proteins in cell culture and *in vivo*. Our hypothesis, is that inclusion of this peptide on the N-termini of delivered proteins will result in increased exogenous protein delivery and an increase in observed luminescence signal.

Additionally, refining ideal concentrations and time segments required to escape endosomes or achieve cytosolic localization of these delivered proteins for optimized luminescence will also be evaluated (concentration and time dependent readouts are expected). Efforts toward these analyses are underway and will be discussed in due time.

6.6 Conclusion – NanoBit

In conclusion, limitations due to the inability of proteinaceous reagents to access the cytoplasm greatly requires the invention of novel delivery methods. This challenge goes in hand with mandatory, straightforward, easy to use new technologies to robustly comment on the apparent success of those methods. As I have presented in this rather large initial data set, intracellular reassembly of NanoBit produces luminescence that appears to be well-behaved and reflective of concentrations of SmBit as a direct result of proteinaceous cellular uptake.

Consistent with Western blot analysis concerning our previously reported polycationic resurfaced nanobodies, we observed robust luminescent signal as an exact outcome of NanoBit intracellular reassembly. Based on this analysis, we believe this platform fulfills the general expectation for measuring cytosolic residence of exogenously delivered proteins. We trust, due to the facile set up this technology, its robust nature and readout, and the use of minimalist tags, that this platform will find great use for the high-throughput screening of proteins and peptides for exogenous protein delivery.

6.7 Part II: Innovate Platforms for *In Vivo* Molecular Tagging and Imaging

A common theme in chemical biology is the promotion of modular chemical attachment to biological reagents, such as proteins and nucleic acids, and whole biological systems. Such control would allow for unprecedented interrogation of such entities and a more complete understanding of disease states and novel routes for therapeutics and diagnostics. Here, we describe the initial investigation of minimalistic and orthogonal molecular assemblies for the visualization of *in vivo* protein and nucleic acid dynamics that readily function in living cells, while experiencing enhanced specificity and kinetics.

As presented throughout this thesis, fluorescent proteins such as GFP have revolutionized chemical biology and fluorescence microscopy. In this chapter, GFP is a key player by its ability to function as a genetically encoded fluorophore, thus selectively lighting up proteins of interest.²⁰⁻
²² However, GFP also has many shortcomings, including its medium size (~25 kDa), slow chromophore formation, and a predisposition towards rapid photobleaching making high-speed or long-term single molecule tracking difficult. Additionally, permanent attachment of such large tags is feared to interfere with underlying biological processes, as well as the current inability to

discriminate or identify post-translational protein modifications once the ‘light’ is on. In response to these disadvantages, researchers have put effort into the manufacture of minimalistic probes and building blocks that can produce fluorescence in a highly-controlled fashion and rely on conceptually different fluorescence tagging schemes.

One such method for intracellular imaging includes the use of cell-permeable biarsenical dyes called Fluorescein Arsenical Hairpin (FlAsH/tetracysteine) and ReAsH (resorufin) reagents.²³⁻²⁴ These platforms utilize a genetically encoded peptide tag (tetracysteine epitope) on the protein of interest that can be recognized with the FlAsH or ReAsH small-molecule reagent. These reagents have been shown useful for visualizing mRNA translation in living cells,²⁵⁻²⁶ but because the reagent is based on arsenic there are concerns of cancer and other arsenic-related diseases as well as cell cytotoxicity.²⁷ Alternatively, antibodies, which allow for specificity and selectivity and can be engineered to recognize and bind virtually any probe or epitope, might appear as ideal reagents for these purposes. However, in addition to their expensive manufacture and large size (~150 kDa), antibodies possess multiple disulfide bonds, making them ill-posed to function in the reducing environment of mammalian cells.²⁸ For these reasons, as presented in Chapter Four, antibody-based mimics including Fabs and scFvs have been investigated as alternatives for *in vivo* imaging (**Figure 6.6A**).²⁹⁻³² Epitopes for which excellent antibodies and thus Fabs and scFvs exist, and include the FLAG³³ and influenza hemmagglutinin (HA)³⁴ tags (**Figure 6.6B**). These and other common epitopes consist roughly of 8-15 residues, however a single copy of such a tag is often insufficient for single molecule tracking due to dimness and background fluorescence. To overcome this limitation, Looger, Viswanathan and co-workers recently engineered GFP to contain numerous (10-12) copies of peptide epitopes (FLAG and HA), termed the spaghetti

monster (SM). This genetically encoded spaghetti monster tag endows simultaneous binding of fluorophore conjugated antibody derivatives as an amplification signal.³⁵

With such molecular assemblies, it is now possible to track proteins in living cells. An example of scFv-based technology for this purpose is the SunTag. It has been used for real-time detection of proteins in living cells and to amplify transcription.³⁶ To visualize proteins, cells co-express a protein that displays many copies (~24+ repeats) of a short linear peptide epitope with concomitant expression of an scFv that recognizes the epitope fused to a fluorescent protein, such as GFP. When the amplification tag is recognized by 24+ scFvs a very bright fluorescent signal results.

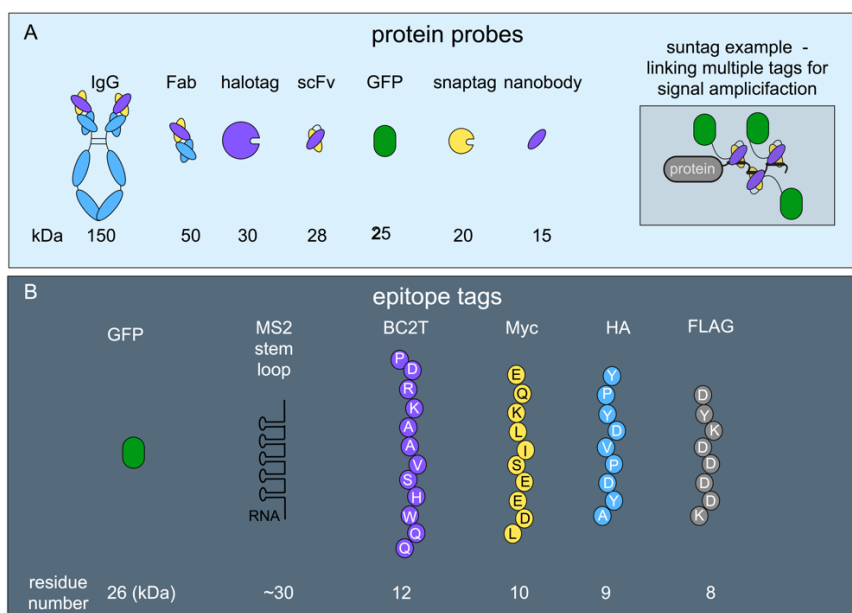


Figure 6. 5 Amplification Tags for Imaging. A) Current options for protein probes. B) Widely used epitope tags. C) Current probe/tag platforms are shaded in the blues. As the size of the probe or tag increases there is generally an increase in specificity and selectivity as well as an increase in probe/tag options. Our goal is to increase the choices of probes and tags that fit within the designated target zones (pink and orange regions).

Current work in the Stasevich Laboratory utilizes a sparsely resolved amplification tags (the previously discussed spaghetti monster in this case) to detect the manufacture of proteins of interest. mRNA encoding a 10X FLAG-SM tagged protein and 24X MS2 tag in the 3' untranslated region is produced in cells. Exogenous delivery of fluorophore conjugated Fabs that recognize FLAG are used to detect proteins of interest following translation of the FLAG tag from the ribosome. Additionally, in order to visualize translation dynamics, the 3' untranslated region of that protein containing the repeat MS2 stem loops (24X) can be detected by an orthogonal fluorophore conjugated MS2 coat protein (**Figure 6.6**).³⁷⁻³⁸ Collectively, co-localization of these two orthogonal fluorophores provide a glimpse into translation dynamics in living mammalian cells.

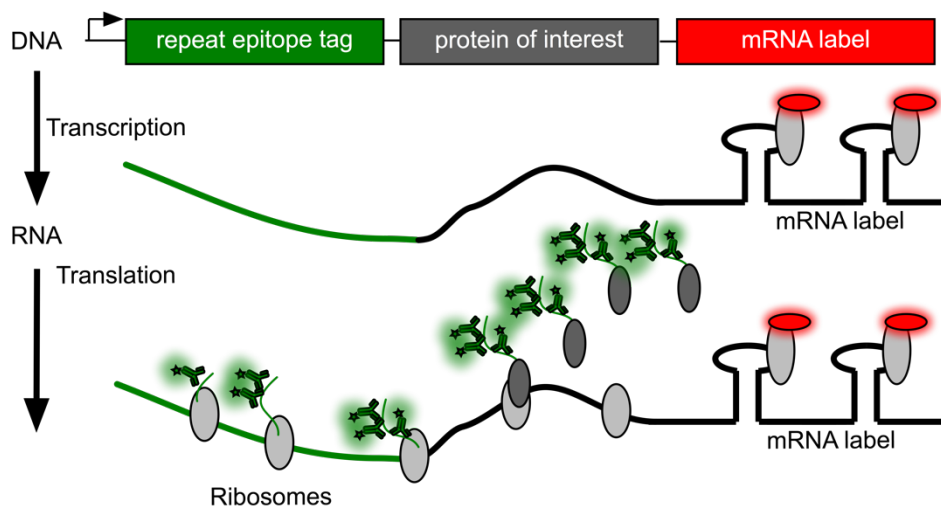


Figure 6. 6 Imaging Translation Dynamics in Living Cells. DNA encoding a protein of interest is flanked with 1) a proteinaceous repeat epitope tag on the N-term for visualization of the protein and 2) a repeat mRNA stem-loop in the 3' untranslated region for mRNA visualization. Typically, scFvs or Fabs are used for identifying the N-term epitope with MCP and PCP recognition on the 3' end. Co-localization of both entities indicates translation. This figure was adapted from *Trends in Genetics*, **2017**, 33, 322.

However, Fabs, like the antibodies they are derived from, are expensive and scFvs that are genetically encoded are often unstable and insoluble in the cytoplasm of cells. This makes finding a good probe that functions *in vivo* analogous to finding a robust antibody as for a Western blot, and once one is found, even harder to find two, thereby limiting multiplexing strategies for the visualization of multiple proteins and events.

Our goal is to develop a new class of protein tags and probes which are minimalistic and orthogonal to natural biology to enable a new level of live-cell imaging (**Figure 6.5C**). These probes will bring pre-existing fluorescence to a protein rather than rely on fluorescence genetically encoded as a fusion protein with the protein of interest. Furthermore, these tags possess high specificity, high affinity and minimal interference when labeling their protein of interest. And finally, because the system does not rely on permanent genetically encoded fluorophores, photobleached probes are less detrimental as they can be substituted for unbleached ones and can potentially allow for different binders based on the stage of the protein's lifecycle and displayed post translational modifications. Here, we describe the development of two examples, both utilizing protein reagents as probes to tag (1) mRNA and (2) short peptide sequences.

6.8 Design of RNA Binders for Imaging Dynamics in Living Cells

mRNA imaging in living cells has generally been accomplished through three methods: (1) exogenous delivery of fluorescently labeled RNA, (2) aptamer RNA labels and (3) recognition by RNA binding proteins that have been fluorescently labeled. Perhaps the most straightforward method is direct RNA microinjection. This method requires fluorophore conjugation of exogenous RNA, typically through chemical methods, and its injection into living cells. While this procedure enables live cell RNA imaging of virtually any RNA and the use of bright, stable fluorophores, it

relies on RNA that was not made in the cell. As a result it may lack interactions with regulatory binding proteins or post transcriptional modifications and furthermore may result in aberrant or abnormal behavior of the RNA or the protein it encodes.³⁹

Alternatively, reactive fluorescent dyes that bind specifically to genetically encoded RNA hairpins have been used to visualize RNA. An example of such probe is Spinach and Spinach2.⁴⁰⁻

⁴¹ These reagents function conceptually similar to genetically encoded GFP, but require the delivery of an activatable fluorescent dye. A true advantage of these probes is that they only turn ‘on’ upon binding RNA, greatly decreasing background fluorescence. These reagents have helped identify noncoding RNAs in cells that are now linked to neurodegenerative disease. Recent work in RNA aptamers is focused on the production of additional colors and minimizing sequence requirements. The current aptamer for Spinach recognition requires a 98 nucleotide (nt) RNA segment. RNA-Mango is one such improvement that requires only 39 nts and experiences higher affinity for its aptamer.⁴² Current shortcomings of this technology include the extended RNA nt requirements for recognition and live cell assay optimization.

Finally, protein-RNA interactions have been exploited for their use of visualizing RNA. This procedure has two components: (1) an RNA binding protein typically fused to a fluorescent protein or conjugated with a fluorophore and (2) RNA engineered to contain the RNA protein binding sites. Perhaps the most widely used system includes phage coat proteins (MS2 coat protein (MCP) and PP7 coat protein (PCP)) that recognize short hairpins.⁴³⁻⁴⁴ Work in the Singer Laboratory formulated this platform for the visualization of mRNA in yeast during cell division by placing multiple MS2 binding sites in the 3’ untranslated region of an mRNA of interest.⁴⁵ An advantage of this approach is the MS2 binding sites can be cloned into any mRNA of choice. The facile setup of this platform has yielded its incorporation into a number of systems that investigate mRNA,

including recent work in the Stasevich Laboratory (**Figure 6.6**).³⁷ Other protein – RNA interactions have also been used to monitor RNA. An example is the RNA recognition motif (RRM) U1A.⁴⁶ Human derived U1A selectively binds the U1hpII RNA hairpin with excellent affinity and selectivity ($K_D \sim 10$ nM). Specifically, U1A is attractive because it is smaller than MS2 and recognizes a slightly shorter nucleotide stem-loop with increased affinity (**Table 6.2**). However, since U1A is of mammalian origin, it's RNA target is endogenously expressed in mammalian cells and therefore can only be used as an orthogonal tag in alternative systems such as yeast.

Table 6. 2 Comparison of MS2 RNA binding protein with U1A and TBP6.7.

Attribute	MS2	U1A	TBP6.7
Length of RNA binding domain (aa)	134	95	109
Molecular weight of RNA binding domain	14.5	10.9	12.5
Affinity to cognate site	~39 nM	~10 nM	~0.5 nM
Size of RNA hairpin binding site	19-nt stem-loop	18-nt stem-loop	~31 nt stem-loop

Alternatively, effort in our lab has utilized the N-terminal RNA recognition motif (RRM) U1A as a protein scaffold to generate synthetic RRM that recognize hairpins of therapeutic interest, such as the HIV transactivation response (TAR) element hairpin.⁴⁷ Yeast display high-throughput screening and saturation mutagenesis generated an RRM we've named TAR Binding Protein 6.7 (TBP6.7) that experiences potent and selective recognition of TAR RNA (apparent $K_D \approx 0.5$ nM) (**Figure 6.7**).⁴⁸ Because there is a limited number of alternatives to the MCP and PCP systems that

can function in mammalian cells (thus greatly restricting multiplexing efforts) we wished to investigate if our recently reported TAR RNA binding protein (TBP 6.7) and TAR RNA could be used in place of MCP/MS2 stem-loop.

In order to experience a bright signal well above background, it is typical to use a 24X repeat of the MS2 stem-loop. As one might image, the cloning of a 24X nucleotide sequence repeat in a short segment is quite challenging. A great deal of effort went into the original making of the 24X

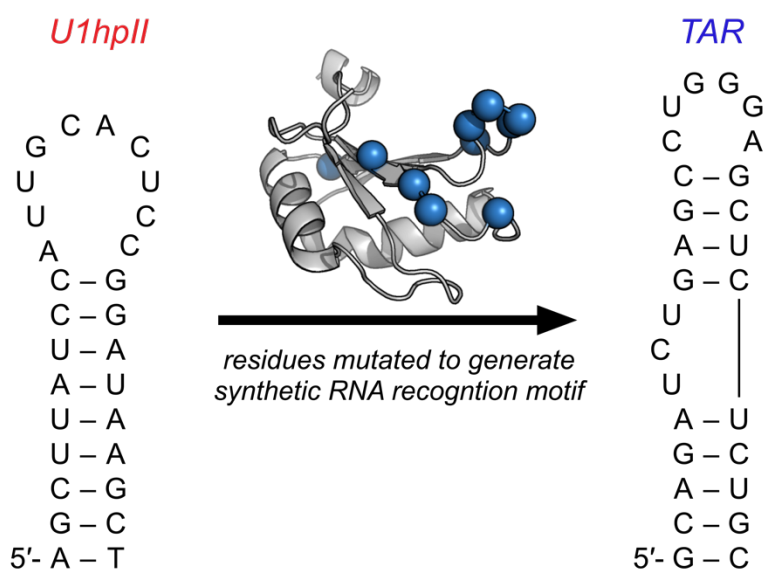


Figure 6. 7 Evolution of U1A protein from U1hpII recognition to recognition of TAR RNA. Residues mutated to generate TAR binding protein are highlighted in blue. PDB: 1URN.

MS2 repeat sequence, and as such they were able to make it with slight variations in the repeats to make cloning possible while still maintaining potent recognition by MCP. However, since our system is a fairly new discovery, we were skeptical of making changes in the stem loop in fear that it might diminish binding. Initial characterization of the TBP 6.7/TAR interaction investigated TAR RNA mutants with slight sequence and secondary structure variations. It was noted that in addition to recognition of TAR RNA, TBP 6.7 also has affinity for a variant of TAR (hp3) where

two guanine groups in the upper hairpin loop are changed to cytosine. To help with cloning, two repeats of the hp3 variant were mixed in with WT TAR RNA as well as two repeats with a modified bulge that can also be recognized by TBP 6.7 (unpublished data) (**Figure 6.8**). This generated a 7X TAR repeat that was readily cloned into a plasmid obtained from the Stasevich Group (spaghetti monster (SM-FLAG) – β -actin – 7XTAR). U2OS mammalian cells were then imaged on the custom built Stasevich Laboratory microscope. Imaging of cells lacking the TAR RNA plasmid, but bead loaded with purified sfGFP-TBP 6.7 (**Figure S6.3**) yielded a similar fluorescence profile and probe distribution as MCP without its cognate RNA (**Figure 6.9**). Additionally, fluorescent recovery after photobleaching (FRAP) of cells bead loaded with sfGFP-TBP6.7 showed quick recovery of fluorescence to that area, indicating minimal endogenous off-target binding.

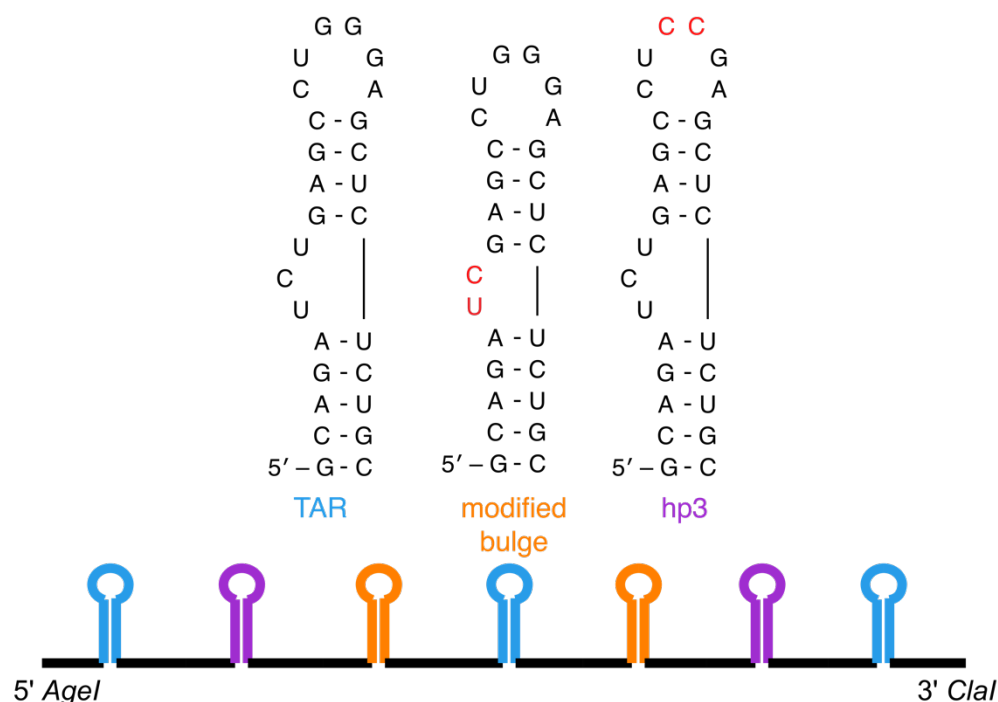


Figure 6. 8 TAR RNA hairpins used to make 7X TAR repeat. WT TAR RNA experienced three repeats, while TAR RNA variants (modified bulge and hp3) experienced two apiece. Insert was flanked with 5' *AgeI* and 3' *ClaI* restriction enzyme sites for vector ligation.

Next, the SM-FLAG – β -actin – 7XTAR plasmid was transfected into U2OS cells. After ~5 hours, cells were washed and sfGFP-TBP 6.7 was bead loaded in along with Fab that recognized the spaghetti monster FLAG construct. **Figure 6.10A** shows β -actin Fab staining as a result of SM-FLAG- β -actin. Staining indicative of β -actin was observed. This suggests that the incorporation of a 7X repeat of TAR RNA in the 3' untranslated region did not interfere with the production of β -actin. Finally, **Figure 6.10B** shows the RNA tagging for imaging this cell using TBP 6.7. To measure the lifetime of sfGFP-TBP6.7 to TAR RNA, we once again performed FRAP experiments in cells transfected with the TAR RNA construct and bead loaded with sfGFP-TBP6.7. We experienced little FRAP recovery in 10 minutes, implying that most sfGFP-TBP6.7 remained bound. This is highly promising for further optimization of this platform, as extended binding kinetics are required for measuring translation elongation times on time scales ranging from ~10 s to ~5 minutes. While, we were unable to find an active translation event (co-localization of Fab and sfGFP-TBP6.7) the RNA binding results were similar to those observed with the MCP/MS2 system, suggesting optimization of delivery and imaging would very likely result in functionality of this system.

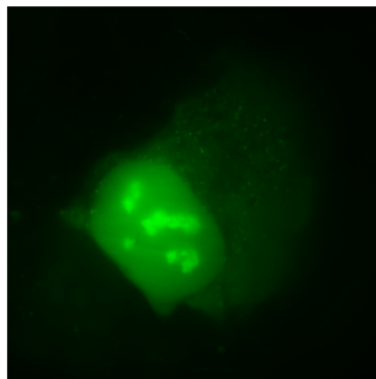


Figure 6. 9 Image of Bead Loaded sfGFP-TBP6.7. Purified sfGFP-TBP6.7 was bead loaded into U2OS cells alone. Images are similar to images of bead loaded MCP-GFP alone.

Use of the MCP/MS2 system benefits from its relatively straightforward design and ability to be implemented as a tag for the visualization of any desired mRNA. Moreover, since it is derived from a virus it experiences orthogonality for imaging in mammalian systems. These same benefits are also apparent in the TAR RNA/TBP 6.7 platform. Additionally, because of similarities in experimental make up, we believe this system will be readily taken up by the general scientific community and will find use as a companion reagent with the MCP/MS2 platform allowing for the multiplexing investigations of RNAs in cells that are currently impossible with current technologies.

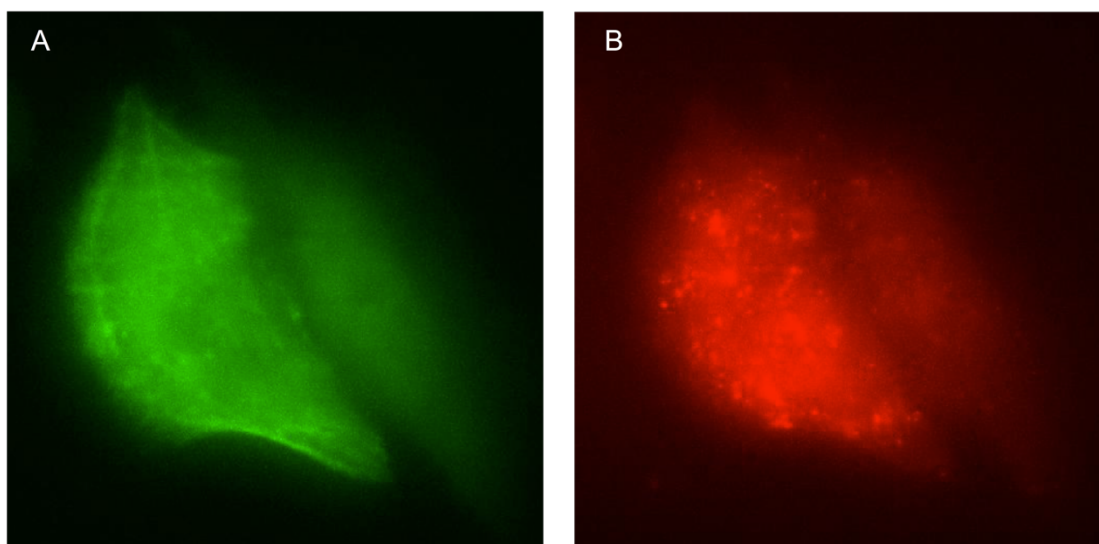


Figure 6. 10 Use of TBP6.7 in Cells Generating TAR RNA Construct. A) Imaging channel for visualizing β -actin using Fabs that recognize SP-FLAG. B) Imaging channel for visualizing TAR RNA using sfGFP-TBP6.7.

6.9 Design of Protein Binders for Imaging Dynamics in Living Cells

Along with the generation of proteins that recognize RNAs, our lab also has experience with the generation of synthetic proteins for protein-protein interactions.⁴⁹⁻⁵⁰ Of particular interest for the purpose of this chapter is our work with helix bundles derived from HIV.^{49, 51} To gain entry into a host cell, binding of HIV-1 gp41 C-peptide interacts with a 5-helix bundle to drag viral and

cell membranes into close proximity, promoting fusion (**Figure 6.11A**).⁵² We hypothesized that replacement of canonical peptide epitopes with C-peptide (which experiences high affinity for the 5-helix bundle), might represent a novel and highly orthogonal interaction for the tagging of proteins (**Figure 6.11B**). To test this concept we cloned a plasmid encoding the C-peptide fused to Histone H2B, a major histone protein for the packaging and maintaining of DNA. Next, we purified sfGFP-5 helix for tagging purposes (**Figure S6.4**). When the H2B fusion protein is expressed in mammalian cells it localizes to the nucleus. We hypothesized that upon delivery of sfGFP-5helix we would see localization of GFP to the nucleus.

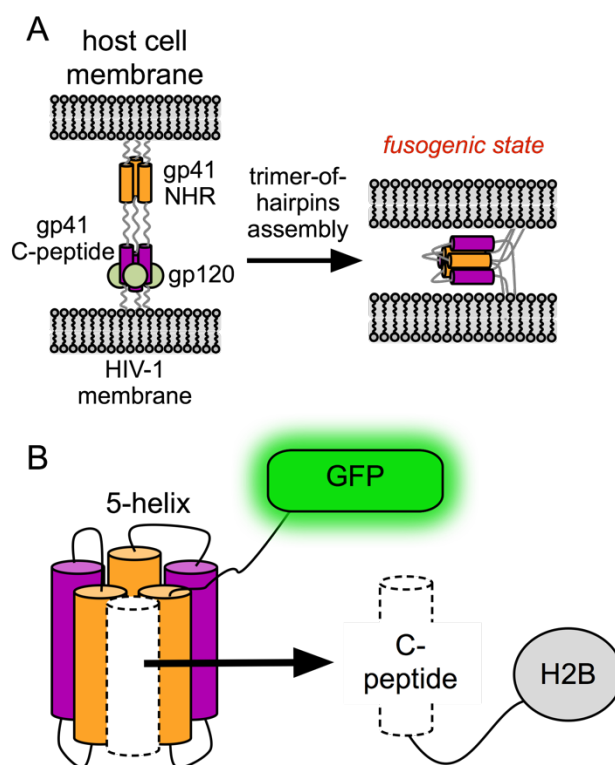


Figure 6. 11 HIV-1 Helix Bundle for Live Cell Imaging. A) The 5-helix/C-peptide platform is derived from HIV-1 gp41-mediated membrane fusion of HIV-1 with a host cell. Binding of a C-peptide (purple) to an N-terminal heptad repeat (NHR, orange) drags viral and cell membranes into close proximity, promoting fusion. B) Recognition of the 5-helix bundle with C-peptide acts as a potential probe/tag platform for non-canonical probes. This figure was adapted from *ChemBioChem*, **2016**, 17, 1945.

Once again, we started with testing the distribution of bead loaded sfGFP-5 helix without expression of its C-peptide. This test experienced slight localization to the nucleus, even without the C-peptide-H2B construct (**Figure 6.12A**). However, FRAP experiments of those cells resulted in quick fluorescence recovery indicating limited off target recognition. Next, U2OS cells were transfected with the plasmid encoding C-peptide-H2B. After a couple of hours, cells were washed and sfGFP-5helix was bead loaded into transfected cells. Imaging of cells showed greater (brighter) localization to the nucleus (this time expected since C-peptide-H2B was present) (**Figure 6.12B-C**). Satisfyingly, FRAP experiments of cells transfected with the C-peptide-H2B construct and bead loaded with sfGFP-5 helix yielded minimal fluorescence recovery well after 10 minutes, also indicating ideal binding kinetics for imaging translation dynamics.

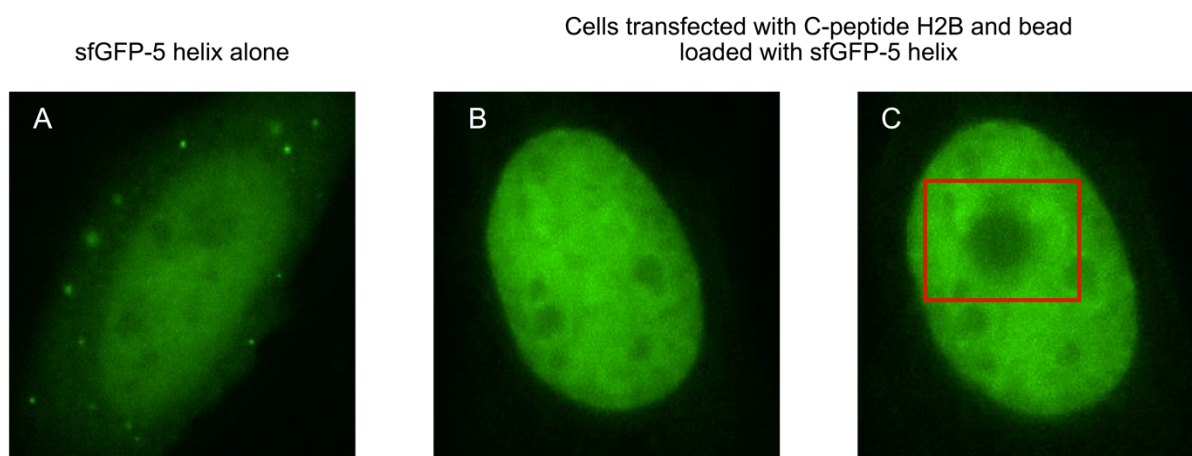


Figure 6. 12 *In Vivo* Cell Imaging with 5-helix/C-Peptide System. A) U2OS cells bead loaded with sfGFP-5 helix alone to test probe distribution within a cell. B) Cells transfected with C-peptide-H2B and then bead loaded with sfGFP-5-helix. Bright fluorescent nuclear localization is observed. C) FRAP experiment to test the binding lifetime of 5-helix/C-peptide. Image is taken 10 minutes after photobleaching area within red box.

As shown in **Figure 6.5A**, common protein probes are highly reliant on antibodies and their derivatives. Additionally, high-throughput methods for the identification of protein based reagents for the recognition of orthogonal epitopes can be generated. At present there is a lack in non-

canonical, creative alternatives for probing mammalian cells that do not rely on antibody derivatives. With the current technologies for protein engineering available, the quest for the generation of minimalistic platforms and technologies for interrogating live cells is more of a question in creativity than possibility. While this analysis is still in its infancy, the case for creative solutions outside the antibody paradigm has been presented and is seemingly supported. Major issues with probes includes challenges of insufficient target specificity, especially in complex *in vivo* settings, insufficient binding dynamics (on/off rates), or insufficient target affinity. I believe answers to these problems can be assessed through the investigation and optimization of probes outside the world of typical antibodies and their derivatives/mimics and the C-peptide/5-helix platform might be one such example.

6.10 Conclusion – Imaging Constructs

There is a great need for orthogonal and efficient molecular assemblies that can label target epitopes while remaining functional in complex and crowded cellular environments. As a proof-of-concept study for the novel use of orthogonal tagging schemes, we show that reagents outside the antibody-based paradigm exist and function comparatively similar (if not better) in complex cellular environments, while drastically reducing complicated manufacture and cost. Further optimization and successful design of these probes will be used to image translational and post-translational gene regulatory dynamics in living cells. In conclusion, such orthogonal and novel tagging technologies that offer unique advantages in size, brightness, intracellular stability, and dynamics can enjoy rapid uptake by the general scientific community. Furthermore, this technology is agnostic to the proposed construct of interest, quantitative where feasible, and has the potential to produce robust dynamics for efficient tagging. We believe there will be great

interest in the technology development and use of these orthogonally engineered and minimalist molecular assemblies.

REFERENCES

1. Fuchs, S. M.; Rutkoski, T. J.; Kung, V. M.; Groeschl, R. T.; Raines, R. T., Increasing the potency of a cytotoxin with an arginine graft. *Protein Eng Des Sel* **2007**, *20* (10), 505-9.
2. Wadia, J. S.; Stan, R. V.; Dowdy, S. F., Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med* **2004**, *10* (3), 310-5.
3. Guo, F.; Gopaul, D. N.; van Duyne, G. D., Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **1997**, *389* (6646), 40-6.
4. Holub, J. M.; Larochelle, J. R.; Appelbaum, J. S.; Schepartz, A., Improved assays for determining the cytosolic access of peptides, proteins, and their mimetics. *Biochemistry* **2013**, *52* (50), 9036-46.
5. LaRochelle, J. R.; Cobb, G. B.; Steinauer, A.; Rhoades, E.; Schepartz, A., Fluorescence correlation spectroscopy reveals highly efficient cytosolic delivery of certain penta-arg proteins and stapled peptides. *J Am Chem Soc* **2015**, *137* (7), 2536-41.
6. Frankel, A. D.; Pabo, C. O., Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988**, *55* (6), 1189-93.
7. Green, M.; Loewenstein, P. M., Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **1988**, *55* (6), 1179-88.
8. Lonn, P.; Kacsinta, A. D.; Cui, X. S.; Hamil, A. S.; Kaulich, M.; Gogoi, K.; Dowdy, S. F., Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Sci Rep* **2016**, *6*, 32301.
9. Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.;

Wood, K. V., NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chem. Biol.* **2016**, *11* (2), 400-408.

10. Bruce, V. J.; Lopez-Islas, M.; McNaughton, B. R., Resurfaced cell-penetrating nanobodies: A potentially general scaffold for intracellularly targeted protein discovery. *Protein Sci* **2016**, *25* (6), 1129-37.

11. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013**, *82*, 775-97.

12. Fuchs, S. M.; Raines, R. T., Arginine grafting to endow cell permeability. *ACS Chem Biol* **2007**, *2* (3), 167-70.

13. Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R., Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* **2010**, *5* (8), 747-52.

14. DePorter, S. M.; Lui, I.; Mohan, U.; McNaughton, B. R., A protein transduction domain with cell uptake and selectivity profiles that are controlled by multivalency effects. *Chem Biol* **2013**, *20* (3), 434-44.

15. DePorter, S. M.; Lui, I.; Bruce, V. J.; Gray, M. A.; Lopez-Islas, M.; McNaughton, B. R., Mutagenesis modulates the uptake efficiency, cell-selectivity, and functional enzyme delivery of a protein transduction domain. *Mol Biosyst* **2014**, *10* (1), 18-23.

16. Thompson, D. B.; Cronican, J. J.; Liu, D. R., Engineering and identifying supercharged proteins for macromolecule delivery into mammalian cells. *Methods Enzymol* **2012**, *503*, 293-319.

17. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R., A class of human proteins that

- deliver functional proteins into mammalian cells in vitro and in vivo. *Chem Biol* **2011**, *18* (7), 833-8.
18. McNaughton, B. R.; Cronican, J. J.; Thompson, D. B.; Liu, D. R., Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. *Proc Natl Acad Sci U S A* **2009**, *106* (15), 6111-6.
19. Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and characterization of a peptide that enhances endosomal escape of delivered proteins in vitro and in vivo. *J Am Chem Soc* **2015**, *137* (44), 14084-93.
20. Shimomura, O., The discovery of aequorin and green fluorescent protein. *J Microsc* **2005**, *217* (Pt 1), 1-15.
21. Tsien, R. Y., The green fluorescent protein. *Annu Rev Biochem* **1998**, *67*, 509-44.
22. Chalfie, M., GFP: Lighting up life. *Proc Natl Acad Sci U S A* **2009**, *106* (25), 10073-80.
23. Machleidt, T.; Robers, M.; Hanson, G. T., Protein labeling with FAsH and ReAsH. *Methods Mol Biol* **2007**, *356*, 209-20.
24. Griffin, B. A.; Adams, S. R.; Tsien, R. Y., Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **1998**, *281* (5374), 269-72.
25. Rodriguez, A. J.; Shenoy, S. M.; Singer, R. H.; Condeelis, J., Visualization of mRNA translation in living cells. *J Cell Biol* **2006**, *175* (1), 67-76.
26. Hoffmann, C.; Gaietta, G.; Zurn, A.; Adams, S. R.; Terrillon, S.; Ellisman, M. H.; Tsien, R. Y.; Lohse, M. J., Fluorescent labeling of tetracysteine-tagged proteins in intact cells. *Nat Protoc* **2010**, *5* (10), 1666-77.
27. Shen, S.; Li, X. F.; Cullen, W. R.; Weinfeld, M.; Le, X. C., Arsenic binding to proteins. *Chem Rev* **2013**, *113* (10), 7769-92.

28. Bruce, V. J.; Ta, A. N.; McNaughton, B. R., Minimalist Antibodies and Mimetics: An Update and Recent Applications. *Chembiochem* **2016**, *17* (20), 1892-1899.
29. Wu, B.; Eliscovich, C.; Yoon, Y. J.; Singer, R. H., Translation dynamics of single mRNAs in live cells and neurons. *Science* **2016**, *352* (6292), 1430-5.
30. Yan, X.; Hoek, T. A.; Vale, R. D.; Tanenbaum, M. E., Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* **2016**, *165* (4), 976-89.
31. Wang, C.; Han, B.; Zhou, R.; Zhuang, X., Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* **2016**, *165* (4), 990-1001.
32. Pichon, X.; Bastide, A.; Safieddine, A.; Chouaib, R.; Samacoits, A.; Basyuk, E.; Peter, M.; Mueller, F.; Bertrand, E., Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *J Cell Biol* **2016**, *214* (6), 769-81.
33. Einhauer, A.; Jungbauer, A., The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. *J Biochem Biophys Methods* **2001**, *49* (1-3), 455-65.
34. Wilson, I. A.; Niman, H. L.; Houghten, R. A.; Cherenson, A. R.; Connolly, M. L.; Lerner, R. A., The structure of an antigenic determinant in a protein. *Cell* **1984**, *37* (3), 767-78.
35. Viswanathan, S.; Williams, M. E.; Bloss, E. B.; Stasevich, T. J.; Speer, C. M.; Nern, A.; Pfeiffer, B. D.; Hooks, B. M.; Li, W. P.; English, B. P.; Tian, T.; Henry, G. L.; Macklin, J. J.; Patel, R.; Gerfen, C. R.; Zhuang, X.; Wang, Y.; Rubin, G. M.; Looger, L. L., High-performance probes for light and electron microscopy. *Nat Methods* **2015**, *12* (6), 568-76.

36. Tanenbaum, M. E.; Gilbert, L. A.; Qi, L. S.; Weissman, J. S.; Vale, R. D., A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **2014**, *159* (3), 635-46.
37. Morisaki, T.; Lyon, K.; DeLuca, K. F.; DeLuca, J. G.; English, B. P.; Zhang, Z.; Lavis, L. D.; Grimm, J. B.; Viswanathan, S.; Looger, L. L.; Lionnet, T.; Stasevich, T. J., Real-time quantification of single RNA translation dynamics in living cells. *Science* **2016**, *352* (6292), 1425-9.
38. Lyon, K.; Stasevich, T. J., Imaging Translational and Post-Translational Gene Regulatory Dynamics in Living Cells with Antibody-Based Probes. *Trends Genet* **2017**, *33* (5), 322-335.
39. Barbarese, E.; Ifrim, M. F.; Hsieh, L.; Guo, C.; Tataavarty, V.; Maggipinto, M. J.; Korza, G.; Tutolo, J. W.; Giampetruzzi, A.; Le, H.; Ma, X. M.; Levine, E.; Bishop, B.; Kim, D. O.; Kuwada, S.; Carson, J. H., Conditional knockout of tumor overexpressed gene in mouse neurons affects RNA granule assembly, granule translation, LTP and short term habituation. *PLoS One* **2013**, *8* (8), e69989.
40. Strack, R. L.; Disney, M. D.; Jaffrey, S. R., A superfolder Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nature Methods* **2013**, *10* (12), 1219-+.
41. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R., RNA Mimics of Green Fluorescent Protein. *Science* **2011**, *333* (6042), 642-646.
42. Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocaru, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J., RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chem. Biol.* **2014**, *9* (10), 2412-2420.

43. Chao, J. A.; Patskovsky, Y.; Almo, S. C.; Singer, R. H., Structural basis for the coevolution of a viral RNA-protein complex. *Nature Structural & Molecular Biology* **2008**, *15* (1), 103-105.
44. Lange, S.; Katayama, Y.; Schmid, M.; Burkacky, O.; Brauchle, C.; Lamb, D. C.; Jansen, R. P., Simultaneous transport of different localized mRNA species revealed by live-cell imaging. *Traffic* **2008**, *9* (8), 1256-1267.
45. Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. H.; Long, R. M., Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **1998**, *2* (4), 437-445.
46. Brodsky, A. S.; Silver, P. A., Identifying proteins that affect mRNA localization in living cells. *Methods* **2002**, *26* (2), 151-5.
47. Blakeley, B. D.; McNaughton, B. R., Synthetic RNA Recognition Motifs That Selectively Recognize HIV-1 Trans-Activation Response Element Hairpin RNA. *ACS Chem. Biol.* **2014**, *9* (6), 1320-1329.
48. Crawford, D. W.; Blakeley, B. D.; Chen, P. H.; Sherpa, C.; Le Grice, S. F.; Laird-Offringa, I. A.; McNaughton, B. R., An Evolved RNA Recognition Motif That Suppresses HIV-1 Tat/TAR-Dependent Transcription. *ACS Chem Biol* **2016**, *11* (8), 2206-15.
49. Walker, S. N.; Tennyson, R. L.; Chapman, A. M.; Kennan, A. J.; McNaughton, B. R., GLUE That Sticks to HIV: A Helix-Grafted GLUE Protein That Selectively Binds the HIV gp41 N-Terminal Helical Region. *Chembiochem* **2015**, *16* (2), 219-222.
50. Chapman, A. M.; McNaughton, B. R., Resurfaced Shape Complementary Proteins That Selectively Bind the Oncoprotein Gankyrin. *ACS Chem. Biol.* **2014**, *9* (10), 2223-2228.
51. Tennyson, R. L.; Walker, S. N.; Ikeda, T.; Harris, R. S.; Kennan, A. J.; McNaughton, B. R., Helix-Grafted Pleckstrin Homology Domains Suppress HIV-1 Infection of CD4-Positive Cells. *Chembiochem* **2016**, *17* (20), 1945-1950.

52. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S., Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **1997**, *89* (2), 263-273.

APPENDICES

Appendix A: Supplemental Information of Chapter Two

Appendix A.1: Chapter Two – Methods

Materials

Phosphate buffered saline (PBS) – Hyclone/Thermo Scientific

0.25% Trypsin – Hyclone / Thermo Scientific

Brilliant Blue R-250 – J.T. Baker

Bovine Serum Albumin (BSA) – Sigma Aldrich

Fetal Bovine Serum (FBS) – PAA Laboratories

Triton X-100 – Fisher Scientific

Dulbecco's Modified Eagle Medium (DMEM) – Hyclone / Thermo Scientific

F-12K Nutrient Mixture (Kaighn's Mod.) – Cellgro / Corning

RPMT-1640 media – Hyclone / Thermo Scientific

Mammalian cell culture dishes – Fisher Scientific

B-PER Bacterial Protein Extraction Reagent – Thermo Scientific

Imidazole – Sigma Aldrich

Modified Lowry Protein Assay Kit – Pierce / Thermo Scientific

Nano-Glo ® Luciferase Assay – Promega

TACS MTT reagent – Trevigen

PageRuler Prestained Protein Ladder – Thermo Scientific

All water was obtained from a Milli-Q water purification system

Instrumentation

All flow cytometry data was carried out on a MoFlo Flow Cytometer and High Speed Cell Sorter with a solid-state iCyt 488 nm laser.

Relative luminescence units were measured on a Synergy Mx Microplate Reader from BioTek

MTT assay absorbance was measured on a Synergy Mx Microplate Reader from BioTek.

Fluorescence microscopy images were taken with EVOS FL from Advanced Microscopy Group.

Mammalian Cell Culture

Human prostate adenocarcinoma cells were cultured in F-12K with 10% Fetal Bovine Serum (FBS) and HEK-293 cells cultured in high glucose Dulbecco's modified Eagle medium (DMEM) with 10% FBS. All cells were incubated at 37 °C with 5% CO₂ environment. All cells were obtained from ATCC.

Cloning

All plasmids were constructed on a pETDuet-1 backbone. All peptides and GGS linkers on the N-terminus and C-terminus of sfGFP were assembled from a set overlapping oligonucleotides. The peptides were then amplified with the sfGFP or nLuc proteins and the constructs were ligated into *NcoI* and *KpnI* restriction enzyme cleavage sites in the pETDuet-1 plasmid.

Protein Purification

BL21 *E. coli* were grown in 500 mL LB cultures at 37 °C to OD₆₀₀ ≈0.6 and induced with 1 mM IPTG at 30 °C overnight. Cells were then collected by centrifugation and stored at -20 °C. Frozen pellets were thawed and 20 mL B-PER was added to lyse cells. The lysate was cleared by centrifugation (17,000 rpm, 30 min) and the supernatant was mixed with 1 mL Ni-NTA agarose resin for 1 hour. The resin was collected by centrifugation (4750, 10 min). The resin was washed with 50 mL of phosphate buffered saline (PBS) with 300 nM NaCl and 20 mM imidazole. The protein was then eluted with 5 mL PBS containing 300 mM NaCl and 500 mM imidazole. The proteins were dialyzed against PBS and analyzed for purity by SDS-PAGE staining with Coomassie Blue. The proteins were then quantified using a modified Lowry protein assay kit. nLuc proteins were purified in the same way, except with Tris buffers (25 mM Tris-HCl, 100 mM NaCl, pH = 8.0) instead of phosphate buffers.

Flow cytometry analysis

Mammalian cells were grown to 90% confluency in a 12-well plate. Cells were then washed once with PBS and 500 µL of diluted protein in PBS was added. The cells were incubated with the protein solution for three hours at 37 °C, 5% CO₂ environment. After the incubation period, cells were then washed once with PBS and two times with PBS-HS (heparin sulfate 20 U mL⁻¹) for 10 minutes at 37 °C, 5% CO₂. The cells were then removed from the dish with 0.5 mL of 0.25% trypsin and collected by centrifugation. The cells were then resuspended in PBS-HS and taken for flow cytometry analysis.

Live Cell Fluorescence Microscopy

Mammalian cells were grown to 90% confluency in a 12-well plate. Cells were then washed once with PBS and 500 μ L of diluted protein in PBS was added. The cells were incubated with the protein solution for three hours at 37 °C, 5% CO₂ environment. After the incubation period, cells were then washed once with PBS and three times with PBS-HS (heparin sulfate 20 U mL⁻¹) for 10 minutes at 37 °C, 5% CO₂. The cells were then imaged on the EVOS FL fluorescence microscope. For 4 °C experiments, the PC-3 cells were incubated at 4 °C for 30 minutes prior to the addition of diluted protein. The incubation period was carried out at 4 °C and washed as described above.

MTT Assay

PC-3 cells were grown to 90% confluency in a 12-well plate. Cells were then washed once with PBS and incubated with the protein in PBS for three hours at 37 °C, 5% CO₂. The solution was removed and the cells were washed twice with PBS-HS (heparin sulfate 20 U mL⁻¹). The cells were then incubated with 0.5 mL medium with 25 μ L of MTT reagent for 4.5 hours. After the incubation, 250 μ L detergent was added to the cells and they were incubated for an additional 30 minutes. MTT assay readings were taken with a Synergy Mx microplate reader at 570 nm.

NanoGlo Luciferase Assay

PC-3 cells were grown to ~80% confluency in a 24-well plate (clear bottom, black well). The nLuc proteins were diluted in TBS (25 mM Tris-HCL, 150 mM NaCl, pH = 7.0) and added to the PC-3 cells. Cells were incubated with each solution for three hours at 37 °C, 5% CO₂. The cells were then washed with TBS, TBS-0.1% tween-20, and TBS-HS (heparin sulfate 20U mL⁻¹). This washing procedure was repeated a total of two times. Then, the cells were incubated with 200 μ L

TBS and 200 uL Nano-Glo Luciferase Assay Reagent for 10 minutes. Luminescence was measured on a Synergy Mx microplate reader.

Appendix A.2: Chapter Two – Proteins Used in this Work

Ypep-GFP

MGYTFGLKTSFNVQGGSGGSGGSGGSMGGASKGEELFTGVVPILVELDGDVNGHKF
SVRGE GEGDATNGKLTLKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKQHDFFKS
AMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNF
NSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQ
SALSKDPNEKRDHMLLEFV TAARITHGMDELYKHHHHHH

Tat-GFP

MGYGRKKRRQRRRGSGGSGGSGGSMGGASKGEELFTGVVPILVELDGDVNGHKF
SVRGE GEGDATNGKLTLKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKQHDFFKS
AMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNF
NSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTP
SALSKDPNEKRDHMLLEFVTAAGITHGMDELYKHHHHHH

Pen-GFP

MGRQIKIWFQNRRMKWKKGGSGGSGGSGGSMGGASKGEELFTGVVPILVELDGDV
NGHKFSVRGE GEGDATNGKLTLKFICTTGKLPVPWPTLVTTLT YGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHK

LEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKHHHHHH

nLuc

MVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSVTPIQRIVLSGENGLKID
IHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHYGTLVIDGVTPNMIDYFGRPYEG
IAVFDGKKITVT GTLWNGNKIIDERLINPDGSLLFRVTINGVTGWRLCERILAHHHHHH

Ypep-nLuc

YTFGLKTSFNVQGG SALALGMVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNL
GVSVTPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVPVDDHHFKVILHY
GTLVIDGVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRVTI
NGVTGWRLCERILAHHHHHH

Appendix A.3: Chapter Two – Supplemental Information

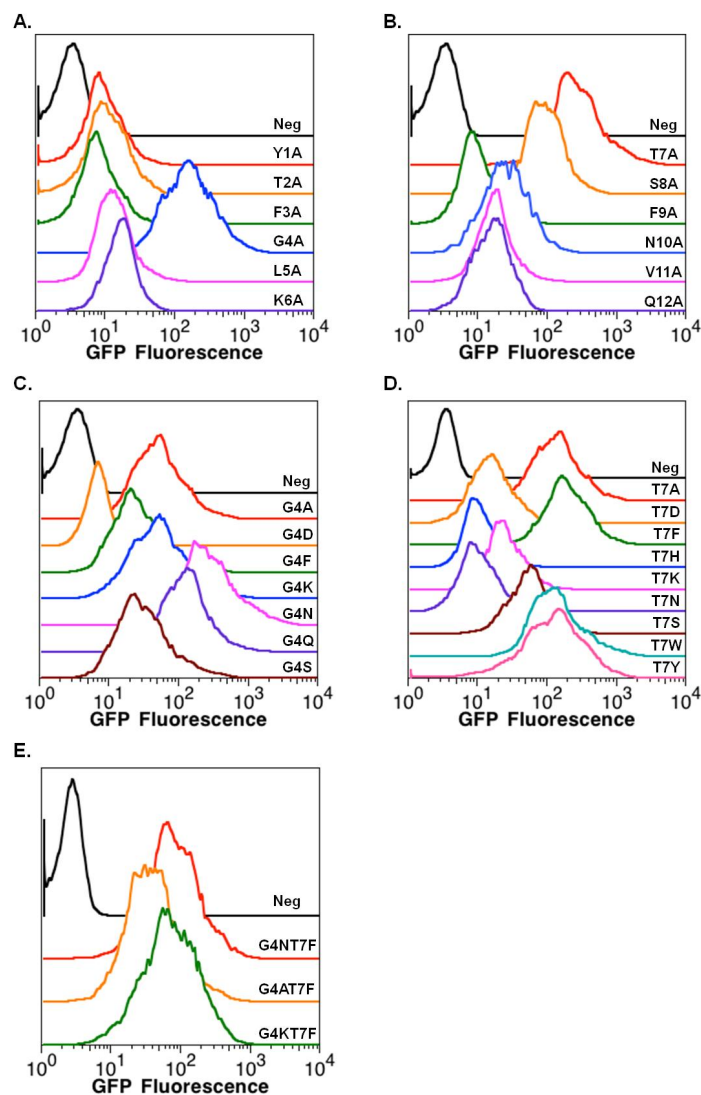


Figure S2. 1 Representative flow cytometry data from Figure 2.2. Flow cytometry data showing (A-B) GFP uptake for alanine mutants of Ypep-GFP. (C) GFP uptake for Ypep-GFP mutants at residue 4. (D) GFP uptake of Ypep-GFP mutants at residue 7. (E) GFP uptake of Ypep-GFP double mutants at residue 4 and 7. (A-E) PC-3 cells treated with 5 μ M mutant Ypep-GFP for 3 hours at 37 $^{\circ}$ C, then washed as described in Methods.

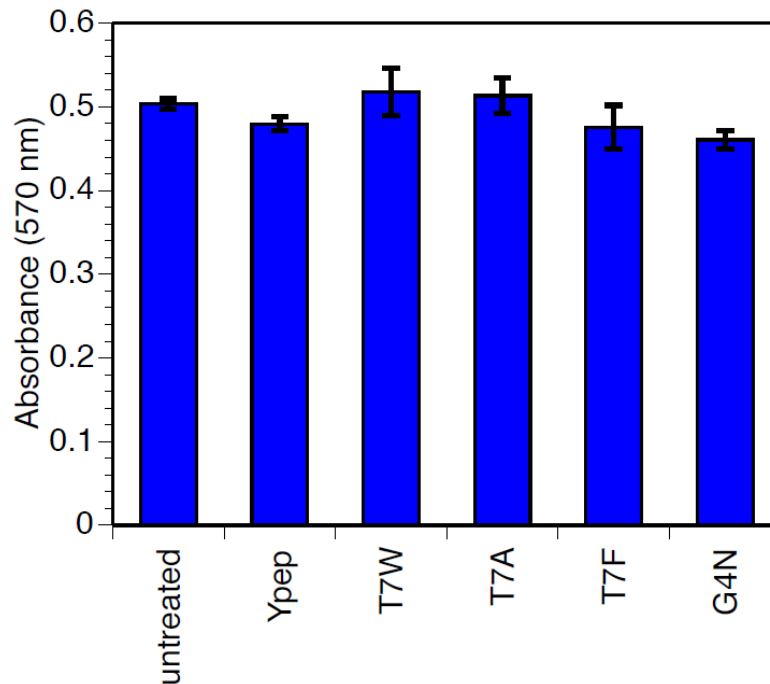


Figure S2. 2 MTT cell viability assay data. PC-3 cells in 12-well plates were treated with 5 μ M protein in PBS or untreated in PBS alone (untreated sample) for 3 hours at 37 $^{\circ}$ C with 5 % CO_2 . The cells were then washed with PBS-heparin sulfate twice, followed by one wash with PBS. The MTT assay was then performed on the cells and the absorbance was read on Synergy Mx Microplate Reader from BioTek. Error bars represent the standard deviation from three independent experiments.

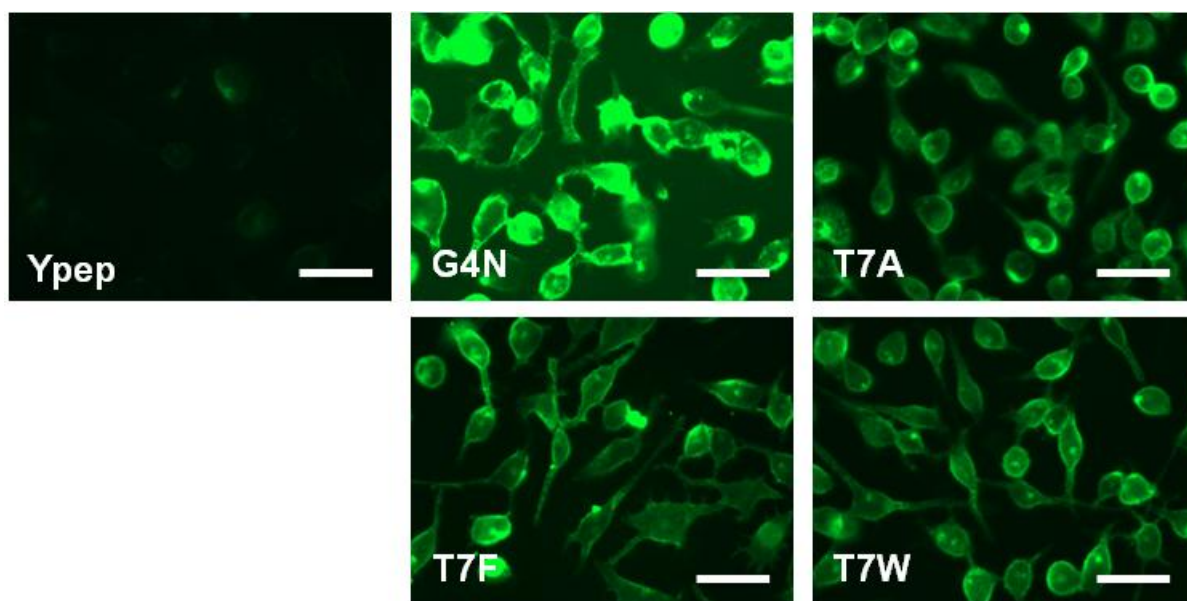


Figure S2. 3 Fluorescence microscopy comparison of Ypep-GFP to mutant Ypep-GFP fusions. Live cell fluorescence microscopy images of PC-3 cells following treatment with 5 μ M of the most efficient mutant Ypep-GFP fusions, then washed to remove cell surface-bound protein. Green color represents internalized GFP. The scale bar is 50 μ M. Lamp intensity was set at 50%, with a 250 ms exposure for all images.

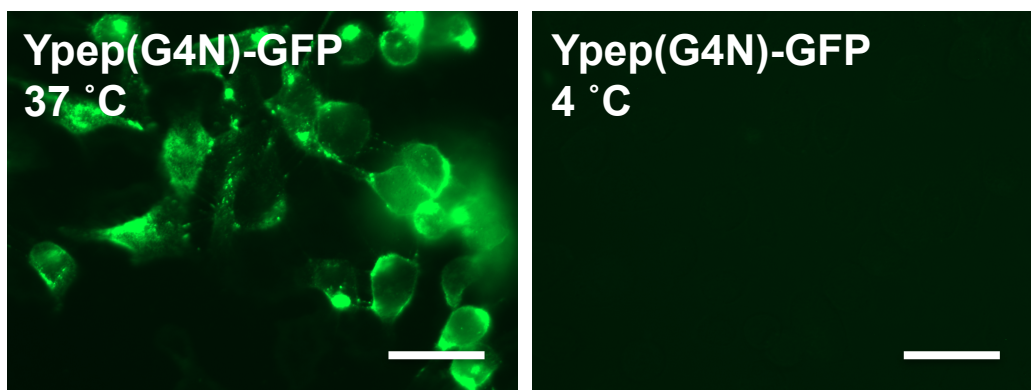


Figure S2. 4 Penetration of G4N-GFP at 37 °C or 4 °C. Live cell fluorescence microscopy images of PC-3 cells following treatment with 5 μ M of (G4N) Ypep-GFP fusions at either 37 °C or 4 °C for 30 minutes. Cells were then washed to remove cell surface-bound protein. Green color represents internalized GFP. The scale bare is 50 μ m. Lamp intensity was set at 50%, with a 250 ms exposure for all images.

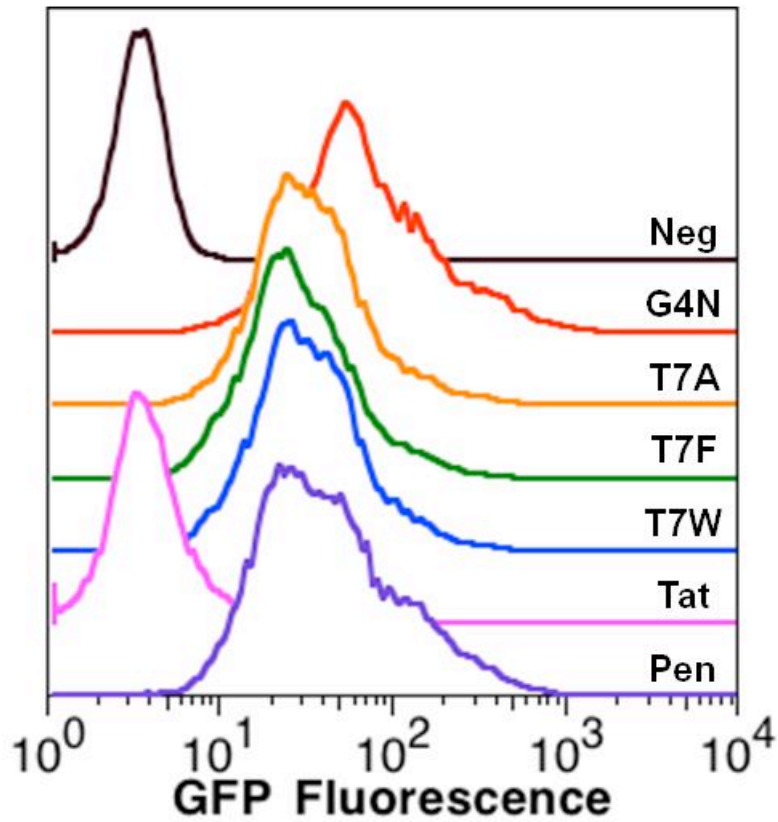


Figure S2. 5 Representative flow cytometry data from Figure 2.4. Flow cytometry data showing GFP internalization of Ypep(mutant)-GFP fusions, Tat-GFP, and penetratin-GFP. PC-3 cells were treated with 1 μ M of each of the proteins for 3 hours at 37 $^{\circ}$ C, then washed as described in Methods.

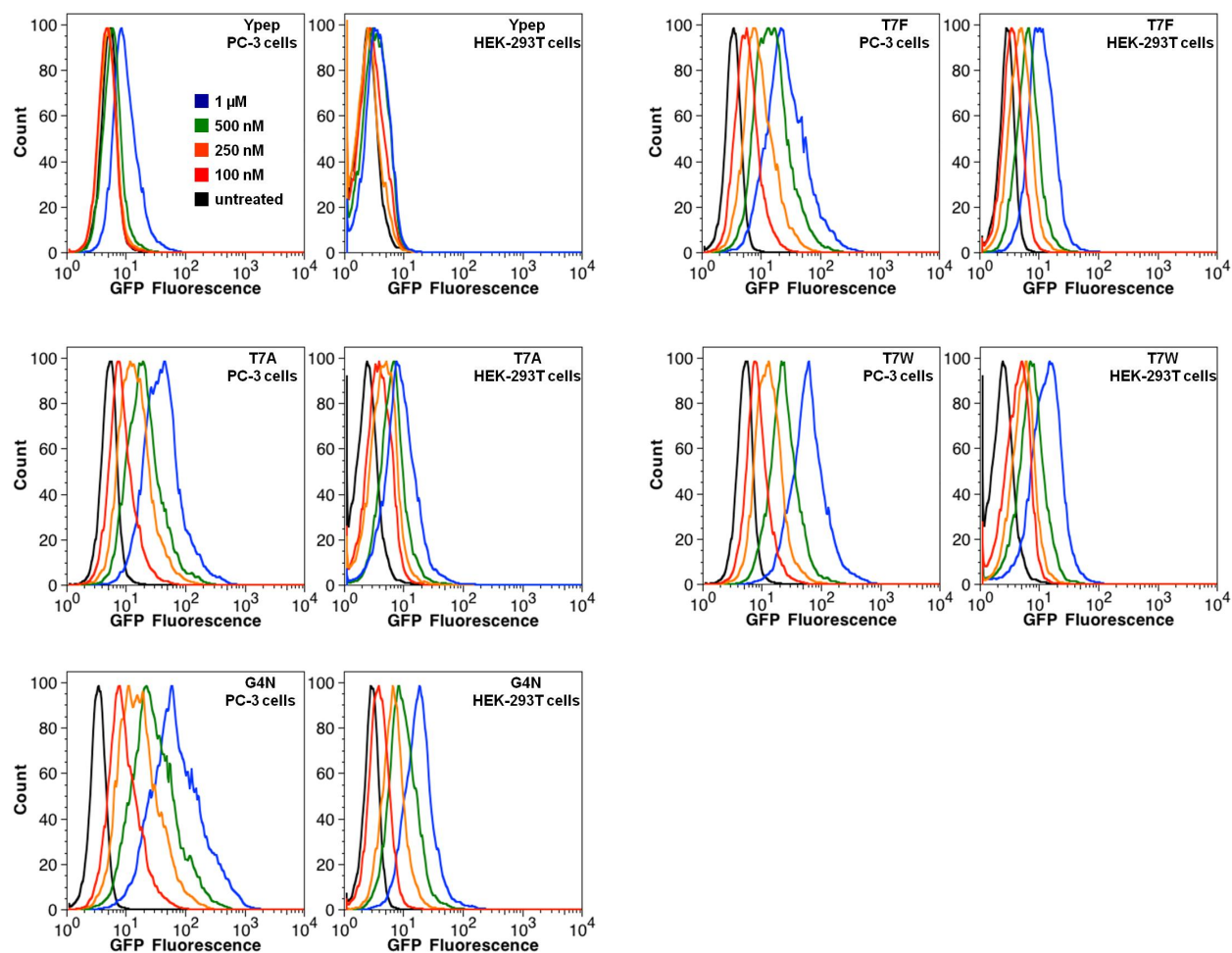


Figure S2. 6 Representative flow cytometry data from Figure 2.5. Flow cytometry data showing the amount of internalized GFP in PC-3 cells or HEK-293 cells following treatment with 100 nM, 250 nM, 500 nM, or 1 μ M Ypep-GFP, Ypep(T7F)-GFP, Ypep(T7A)-GFP, Ypep(T7W)-GFP, or Ypep(G4N)-GFP for 3 hours at 37 $^{\circ}$ C in PBS.

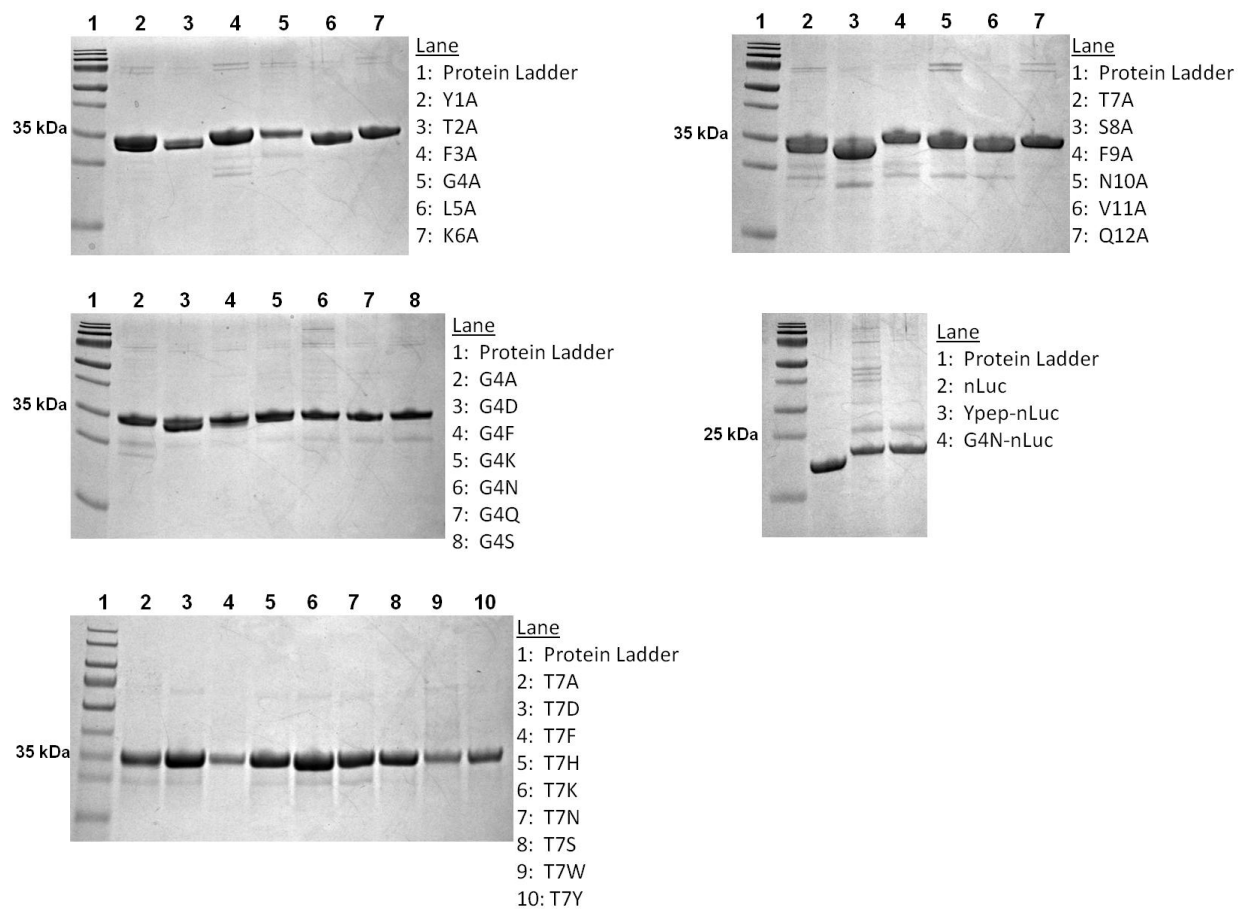


Figure S2. 7 PAGE analysis of all proteins used in this work. Proteins were run on a 15% Tris-HCl gel (BioRad), and stained with Coomassie Blue.

Appendix B: Supplemental Information of Chapter Three

Appendix B.1: Chapter Two – Methods

Materials

All chemicals obtained from Sigma-Aldrich unless specified

LB Miller Broth – Fisher

Phosphate buffered saline (PBS) – Corning Cell Grow

0.25% Trypsin – Hyclone/Thermo Scientific

Fetal Bovine Serum (FBS) – Atlanta Biologicals

Dulbecco's Modified Eagle Medium (DMEM) – Hyclone/Thermo Scientific

Mammalian cell culture dishes – Fisher Scientific

RIPA Buffer – Boston BioProducts

5-alpha chemically competent *E. coli* – NEB

BL21 (DE3) chemically competent *E. coli* – NEB

Agar – GoldBio Technology

Carbenicillin – GoldBio Technology

Restriction Enzymes – NEB

Isopropyl- β -D-1-thiogalactopyranoside (IPTG) – GoldBio Technology

cComplete ULTRA Tablets, Mini, EDTA-free – Roche

Quick Ligation Kit – NEB

Vent Polymerase – NEB

Oligonucleotides – IDT

Miniprep Kits – OMEGA

All antibodies obtained from Abcam unless specified

Rab5 antibody – Cell Signaling Technologies

iBlot Gel Transfer Stack Kit – Novex

PageRuler Prestained Protein Ladder – Thermo Scientific

Ready Gel Precast Gels – BioRad

3T3 Cells –ATCC

All water was obtained from a Milli-Q water purification system.

Instrumentation

MoFlo Flow Cytometer and High Speed Cell Sorter with a solid state iCyt 488 nm laser.

Sonifer W-350 cell disrupter – Branson

Fluorescence microscopy images were taken with EVOS FL from Advanced Microscopy Group

MJ Mini Gradient Thermal Cycler – BioRad

Molecular Imager Gel Doc XR+ System – BioRad

Circular Dichroism Spectrometer – Aviv Model 202

iBlot Apparatus – Invitrogen

Odyssey Classic Infrared Imager – LI-COR

iTC200 – Microcal (Malvern)

Cloning

All plasmids were constructed on a pETDuet-1 backbone. All proteins were assembled from a set of overlapping oligonucleotides. Proteins were amplified using vent and the constructs were ligated into *NcoI* and *NotI* restriction enzyme cleavage sites in the pETDuet-1 plasmid. Proteins containing GFP fusions were assembled from a set of overlapping oligonucleotides and ligated into *NdeI* and *KpnI* restriction enzyme cleavage sites in the pETDuet-1 plasmid.

Protein Purification

Plasmids were transformed into BL21s (DE3). Cells were grown in either 2500 or 500 mL LB cultures containing carbenicillin at 37 °C to an $OD_{600} \approx 0.6$ and induced with 1 mM IPTG at 25 °C overnight. Cells were then collected by centrifugation and resuspended in either phosphate buffer with 150 mM NaCl for NBs (20 mM Sodium Phosphate, pH 7.4) or resuspended in phosphate buffer with 2M NaCl for pcNBs (20 mM Sodium Phosphate, pH 7.4) and stored at -20 °C. Frozen pellets were thawed and incubated with complete ULTRA protease inhibitor tablets then sonicated for 2 min. The lysate was cleared by centrifugation (9000 rpm, 20 min) and the supernatant was mixed with 1 mL Ni-NTA resin for 30 min. The resin was collected by centrifugation (4750 rpm, 10 min). The resin was washed with 50 mL buffer and 20 mM imidazole then 10 mL buffer and 50 mM imidazole. The protein was then eluted with 7 mL buffer containing 300 mM imidazole. The proteins were dialyzed against buffer and analyzed for purity by SDS-PAGE. Purified proteins were quantified using absorbance at 280 nM.

Circular Dichroism

Proteins were purified as described above. Separately, each protein was diluted to 6-8 μM in Sodium Phosphate buffer (20 mM Sodium Phosphate, pH 7.4 and 150 mM NaCl). Wavelength data are the average of three scans from 250 to 200 nm in 1 nm steps at 25 °C.

Mammalian Cell Culture

NIH/3T3 cells were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% Fetal Bovine Serum (FBS). All cells were incubated at 37 °C with 5% CO₂ environment.

Live Cell Fluorescence Microscopy

Mammalian cells were grown to ~80% confluency in a six-well plate. Cells were then washed once with PBS and 2 mL of 250 nM protein fused with GFP was added. The cells were incubated with the protein solution for 3 hours at 37 °C, 5% CO₂ environment. After the incubation period, cells were washed once with PBS and three times with PBS-HS (heparin sulfate 20 U/mL) for 10 min at 37 °C, 5% CO₂. The cells were then imaged on the EVOS FL fluorescence microscope.

Flow Cytometry

Mammalian cells were grown to 80% confluency in a six-well plate. Cells were then washed once with PBS and 2 mL of 10 nM, 250 nM, or 500 nM protein fused with GFP was added. The cells were incubated with the protein solution for 3 hours at 37 °C, 5% CO₂ environment. After the incubation period, cells were washed once with PBS and three times with PBS-HS (heparin sulfate 20U/mL) for 10 min at 37 °C, 5% CO₂. The cells were then removed from dish with 0.25%

trypsin-EDTA and collected by centrifugation. The cells were then suspended in PBS and taken for flow cytometry analysis.

Cytosolic Protein Extraction and Whole Cell Lysate Preparation for Western Blot

3T3 cells were plated in a six-well plate and grown to ~80% confluency. The cells were treated with 250 nM or 500 nM proteins (wtNB-GFP and pcNB-GFP or wtNB and pcNB, respectively) for 24 hours at 37 °C, 5% CO₂. After treatment, cells were washed once with PBS and once with PBS-HS (heparin sulfate 20 U/mL) for 10 min at 37 °C, 5% CO₂ then lifted with 0.25% trypsin-EDTA and pelleted. For cytosolic protein extraction, cell pellets were resuspended in 100 µL of 50 µg mL⁻¹ digitonin in 75 mM NaCl, 1 mM NaH₂PO₄, 8 mM Na₂HPO₄, 250 mM sucrose supplemented with Roche protease inhibitor cocktail tablet and incubated on ice for 10 min. Cells were then centrifuged for 5 min at 13,000 rpm. Supernatant was then used as cytosolic protein extraction. Left over pellets were then re-suspended in 100 µL RIPA buffer supplemented with a Roche protease inhibitor cocktail tablet and incubated on ice for 5 min then further lysed through a 20-gauge needle. Supernatant was then used as whole cell lysate extraction. Both supernatants were collected and separated by SDS-PAGE and transferred to a nitrocellulose membrane via an iBlot Western blotting apparatus. The membrane was incubated with 1X TBS with 5% milk at 25 °C for 1 hour. The membrane was then washed three times with 1X TBS and 0.1% Tween-20 (5 minute incubations). Primary antibodies for GFP, Erk1/2 and Rab5 were incubated with the membrane containing GFP fused nanobodies overnight in 10 mL of 1X TBS, 5% BSA, and 0.1% Tween-20 at 4 °C. The Western blot containing unfused nanobodies were incubated with primary antibodies for His₆, Erk1/2, and Rab5 overnight in the same mixture. Both membranes were washed 3X with 1X TBS containing 0.1% Tween-20 (5 minute incubations) and then incubated

with anti-Rabbit (Alexa Fluor 790) in 10 mL TBS, 5% milk and 0.1% Tween-20 for 1 hour at 25 °C. The membrane was washed 3X with 1X TBS containing 0.1% Tween-20 (5 minute incubations) and imaged in 1X TBS using the Odyssey Classic Infrared Imager.

Ni-NTA Pull Down Assay

wtNB1 and pcNB1 (nanobodies for GFP) tagged with His₆ were cloned into MCS1 of pETDuet-1 using restriction enzymes *NcoI* and *NotI*. Untagged GFP was cloned into MCS2 of pETDuet-1 using restriction enzyme *NdeI* and *KpnI*. Completed constructs were transformed into BL21s (DE3). Cells containing the co-cloned pair were inoculated and induced as described previously. Cells were pelleted and purified as described previously. The pull-down was analyzed by SDS-PAGE.

Appendix B.2: Chapter Three – Proteins Used in this Work

NB1 (GFP NB)

MGMQVQLVESGGALVQPGGSLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAG
MSSAGDRSSYEDSVKGRFTISRDDARNTVYLMNSLKPEDTAVYYCNCNVNMGFEYWGQ
GTQVTVSSHSHHHHHH

pcNB1 (GFP NB)

MQVQLVEKGGKRVQPGGSLRLKCAASGFPVNRYSMRWYRQAPGKEREWVAGMSS
AGDRSSYEDSVKGRFKIKRDDARNTVYLRMRKLKPEDTAVYYCNCNVNMGFEYWGQGTR
VTVSKKHSHHHHHH

NB2 (HER2 NB)

MEVQLVESGGGLVQAGGSLRLSCAASGITFSINTMGWYRQAPGKQRELVALISSIGD
TYYADSVKGRFTISRDNKNTVYLMNSLKPEDTAVYYCKRFRTAAQGTDYWGQGTQ
VTVSSHHHHHHH

pcNB2 (HER2 NB)

MEVQLVEKGGGRVQAGGSLRLRCAASGITFSINTMGWYRQAPGKQRELVALISSIG
DTYYADSVKGRFRIRDNKNTVYLRMRRLKPEDTAVYYCKRFRTAAQGTDYWGQGT
RVTVSKHHHHHHH

NB3 (Beta-Lac NB)

MAQVQLVESGGGSVQAGGSLRLSCTASGGSEYSYSTFSLGWFRQAPGQEREAVAAI
ASMGGLTYYADSVKGRFTISRDNKNTVTLMNNLKPEDTAIYYCAAVRGYFMRLPSS
HNFYWGQGTQVTVSSHHHHHHH

pcNB3 (Beta-Lac NB)

MAQVQLVEKGGGKVRAGGKLRLRCTASGGSEYSYSTFSLGWFRQAPGQEREAVAA
IASMGGLTYYADSVKGRFKIKRDNKNTVTLMNNLKPEDTAIYYCAAVRGYFMRLPS
SHNFYWGQGT RVTVSRHHHHHHH

Appendix B.3: Chapter Three – Supplemental Information

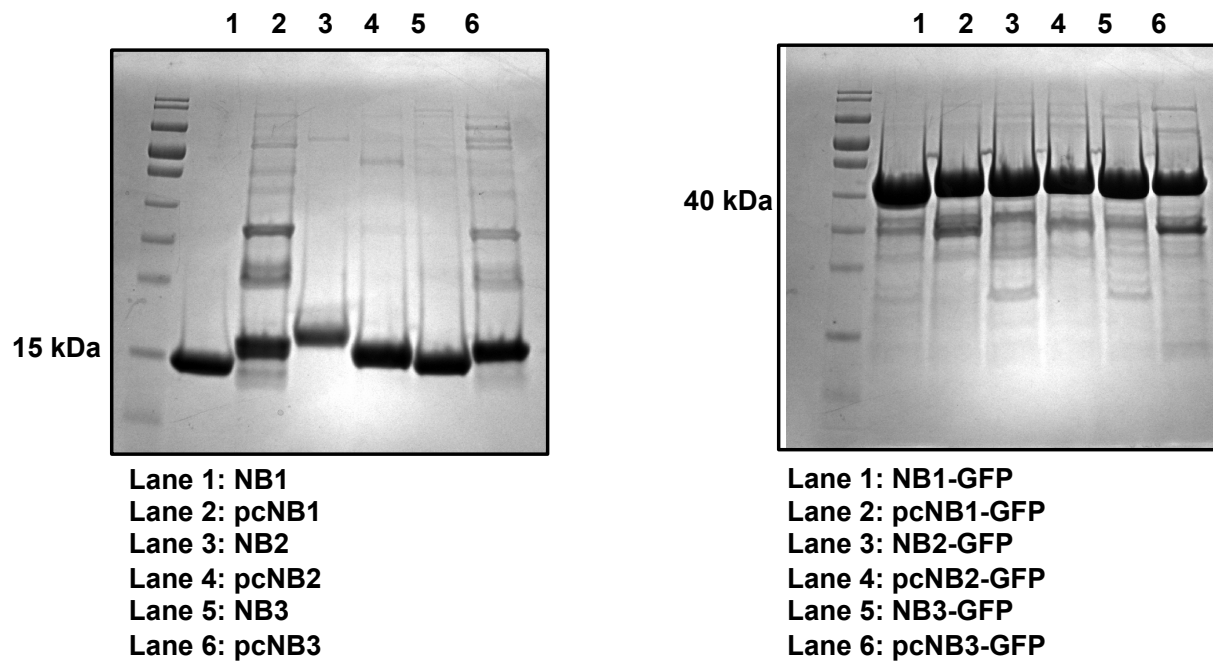


Figure S3. 1 SDS-PAGE of Purified Proteins. PAGE analysis of purified WT nanobody and polycationic resurfaced mutants.

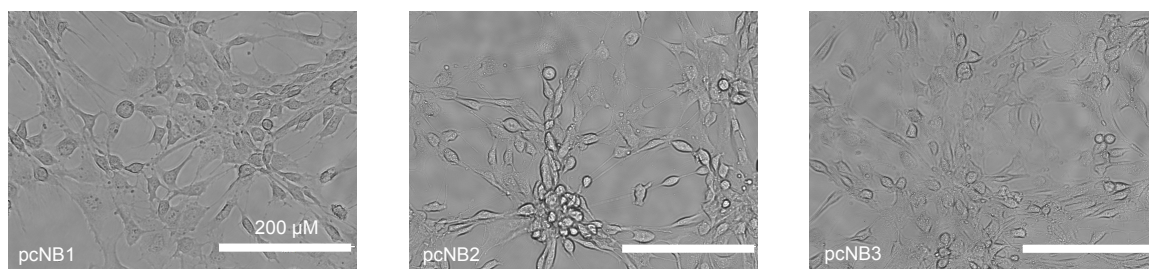


Figure S3. 2 Brightfield Images of Mammalian Cells. Microscopy images of 3T3 cells following treatment with 250 nM resurfaced nanobody-GFP fusions.

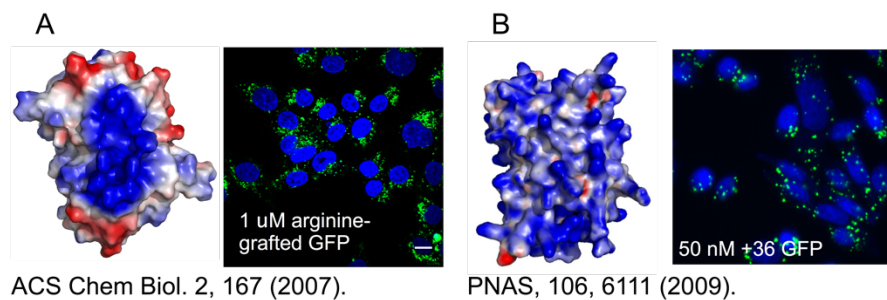
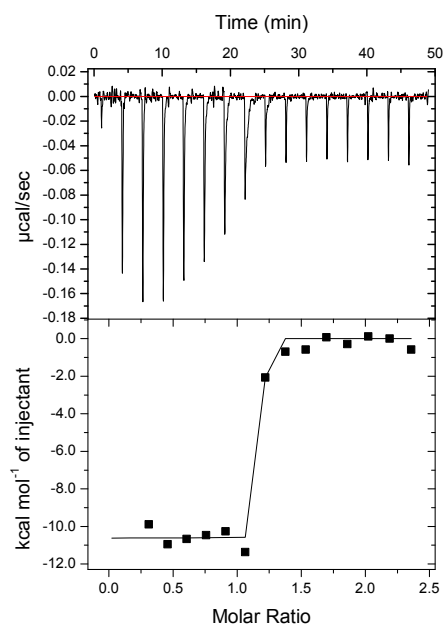
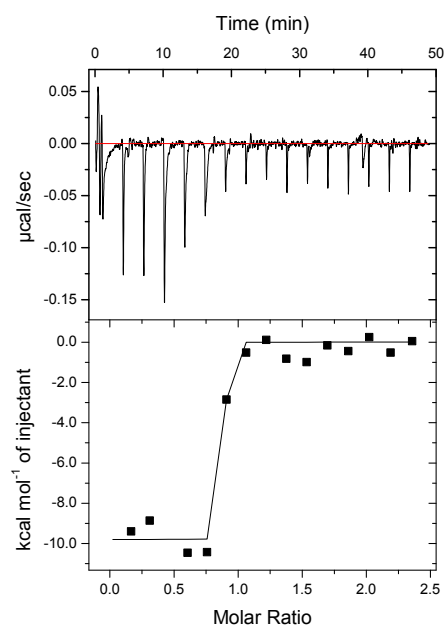


Figure S3. 3 Fluorescent Microscopy Images of Supercharged GFP Variants. GFP (green) appears as punctate foci, suggesting entrapment of proteins in the endosomes.



NB1



pcNB1

Figure S3. 4 ITC Data. Representative ITC binding isotherms involving NB1 and pcNB1 with EGFP.

Appendix C: Supplemental Information of Chapter Five

Appendix C.1: Chapter Five – Methods

Materials

All chemicals were obtained from Sigma-Aldrich unless specified

LB Miller Broth – Fisher

Phosphate Buffered Saline (PBS) – Corning Cell Grow

5 – alpha chemically competent *E. coli* – NEB

BL21 (DE3) chemically competent *E. coli* – NEB

Agarose A – Bio Basic Inc.

Carbenicillin – GoldBio Technology

Restriction Enzymes – NEB

Isopropyl- β -D-1-thiogalactopyranoside (IPTG) – GoldBio Technology

cOmplete ULTRA Tablets, Mini, EDTA-free – Roche

Quick Ligation Kit – NEB

Vent Polymerase – NEB

Oligonucleotides – IDT

Miniprep Kits – OMEGA

All antibodies obtained from Abcam

iBlot gel transfer stack kit – Novex

PageRuler Prestained Protein Ladder – Thermo Scientific

12% Ready Gel Precast Gels – BioRad

Snakeskin Dialysis Tubing 10K MWCO – Thermo Scientific

BirA-Biotin Ligase Kit – Avidity

Casamino Acids – Fisher

Yeast Nitrogen Base without Amino Acids – BD

Yeast Extract – Alfa Aesar

Dextrose – Fisher

D-Galactose – GoldBio Technology

Peptone – Fisher

Penicillin – Streptomycin – Fisher

Gene Pulser/Micro Pulser Cuvettes 1 mm and 2 mm – BioRad

Odyssey Blocking Buffer – LI-COR

TMB One Substrate – Promega

L-(+)-Arabinose – Sigma Aldrich

Chloramphenicol – GoldBio Technologies

Sulfo-Cyanine5 Maleimide – Lumiprobe

IRDye800 CW Maleimide – LI-COR

Albumin, Bovine Fraction V (BSA) – RPI

Streptavidin Coated 96-well Plates (clear and black) – Pierce

EZ-Link Maleimide Activated Horseradish Peroxidase – ThermoFisher Scientific

NanoGlo reagent – Promega

All water obtained from Milli-Q water purification system

Instrumentation

CyAn-ADP flow cytometry cell analyzer

Sonifer W – 350 cell disrupter – Branson

MJ mini gradient thermal cycler – BioRad

Molecular Imager Gel Doc XR+ System –BioRad

iBlot Apparatus – Invitrogen

Odyssey Classic Infrared Imager – LI-COR

Synergy Mx Microplate Reader – BioTek

NanoDrop 200 UV-Vis Spectrophotometer – Thermo Scientific

Cloning

Purified proteins: All plasmids were constructed on a pETDuet-1 backbone. All proteins were assembled from a set of overlapping oligonucleotides or purchased g-block. Constructs were amplified using vent and then ligated into *NcoI* and *KpnI* restriction enzyme cleavage sites in the pETDuet-1 plasmid.

Display vectors: EBY100 yeast (*trp*-, *leu*-, with the *Agalp* gene stably integrated) and pCTCON2 plasmid were generously provided by the Wittrop lab (MIT). The gene coding for mSA2 flagged with C-terminal BC2T were PCR amplified using vent and the constructs were ligated into *NheI* and *BamHI* restriction enzyme cleavage sites in the pCTCON2 plasmid.

MC1061 bacteria electrocompetent cells and pB33-eCPX plasmid were generously provided by the Daugherty lab (UCSB). The gene coding for BC2T-mSA2-myc were PCR amplified using

vent and the constructs were ligated into *NdeI* and *XhoI* restriction enzyme cleavage sites in the pB33-eCPX plasmid.

Protein Purification

Plasmids were transformed into BL21s (DE3). Cells were grown in either 2500 or 500 mL LB cultures containing carbenicillin at 37 °C to OD₆₀₀ ≈ 0.5 and induced with 1 mM IPTG at 20 °C overnight. Cells were then collected by centrifugation and resuspended in phosphate buffer with 2 M NaCl (20 mM Sodium Phosphate, pH 7.4) and stored at -20 °C. Frozen pellets were thawed and incubated with cOmplete ULTRA protease inhibitor tablets then sonicated for 2 minutes. The lysate was cleared by centrifugation (8000 rpm, 20 minutes) and the supernatant was mixed with 1 mL Ni-NTA resin for 30 minutes. The resin was collected by centrifugation (4750 rpm, 10 minutes). The resin was washed with 50 mL buffer and 20 mM imidazole then 10 mL buffer and 50 mM imidazole. The protein was then eluted with 7 mL buffer containing 400 mM imidazole. The proteins were dialyzed against buffer with 150 mM NaCl and analyzed for purity by SDS-PAGE. Purified proteins were quantified using absorbance at 280 nm.

Protein Conjugation

Nanobody Dy Conjugation/HRP: Purified BC2 nanobodies with a C-terminal Cysteine residue were reacted with maleimide dye conjugates or maleimide HRP as described by manufactures' instructions. Briefly, ~10-20-fold molar excess of dye over protein was added to nanobody solution in PBS, mixed and incubated at room temperature for 2 hours to overnight. Final product was separated from unreacted materials via 30,000 MWCO centrifugal columns. The dye labeled

nanobody or HRP labeled nanobody was then purified by dialysis and analyzed via SDS-PAGE. It was stored, protected from light, at 4 °C until ready for use.

Protein – Biotin Conjugation: GFP was conjugated using Avidity BioMix protocols and purified BirA Protein Ligase at 1.0 mg/mL.

ELISA binding assay

HRP: ELISA assays were performed using clear, streptavidin coated, 96-well plates (Pierce). The plate was washed 3 times with wash buffer (20 mM phosphate, 150 mM NaCl, 0.05% Tween-20, and 0.1 mg/mL BSA, pH = 7.4). Following washing, 100 μ L of biotinylated GFP at 10 μ g/mL was incubated for 2 hours at RT. Wells were washed three times with 200 μ L of wash buffer shaking for 5 minutes. Subsequently, wells containing GFP were then incubated for 1 hour at RT with 100 μ L of buffer containing one of three different proteins, all at 50 nM: (1) a BC2 tagged protein that has no appreciable affinity for GFP (zinc finger protein HRX, referred to as HRX); (2) a GFP binding nanobody-His₆ that tightly bind GFP ($K_D \approx 1$ nM), but lacks the BC2T epitope, or (3) GFP binding nanobody fused to a C-terminal BC2T or *myc* tag, then washed three times with 200 μ L wash buffer. Following this, a 1:10,000 dilution of HRP-conjugated anti-His_{6X} or anti-myc antibody were incubated in 100 μ L Odyssey Blocking buffer separately for all samples and ~ 50 nM solution of BC2nb-HRP in 100 μ L Odyssey Blocking Buffer for a separate set of all constructs for 1 hour at RT, and washed 3 times with 200 μ L wash buffer. Colorimetry was developed for 20 minutes using 100 μ L of TMB-One substrate. Absorbance was measured at 655 nm on a plate reader.

NanoLuciferase: ELISA assays were performed using black, streptavidin coated, 96-well plates (Pierce). The plate was washed three times with wash buffer (20 mM phosphate, 150 mM NaCl, 0.05% Tween-20, and 0.1 mg/mL BSA, pH = 7.4). Following washing, 100 μ L of biotinylated GFP at 10 μ g/mL was incubated for 2 hours at RT. Wells were washed three times with 200 μ L wash buffer, shaking for 5 minutes. Subsequently, wells containing GFP were then incubated for 1 hour at RT with 100 μ L of buffer containing one of three proteins, all at 50 nM: (1) a BC2 tagged protein that has no appreciable affinity for GFP (zinc finger protein HRX, referred to as HRX); (2) a GFP binding nanobody-His₆X that tightly binds GFP ($K_D \approx 1$ nM), but lacks the BC2T epitope, or (3) GFP binding nanobody fused to a C-terminal BC2T, then washed three times with 200 μ L wash buffer. Following this, 50 nM of BC2-nanobody-nLuc fusion protein in wash buffer was incubated for 1 hour, and washed four times with 200 μ L wash buffer. Finally 100 μ L of NanoGlo reagent substrate diluted 1:50 in wash buffer was incubated with samples and allowed to shake for ~10 minutes at RT. Luminescence was measured on a plate reader.

Flow Cytometry Analysis

Bacteria: 50 mL culture of bacteria displaying mSA2 with BC2T and myc were grown in a 250 mL baffled flask containing chloramphenicol at 37 °C with shaking (250 rpm) until an OD₆₀₀ = ~0.5 and induced with a final concentration of 0.02% (w/v) L-(+)-Arabinose at 20 °C overnight with shaking (250 rpm). Approximately 10⁸ cells were pelleted and washed with 500 μ L of 4 °C PBS-BSA. Bacteria were subsequently incubated with either BC2 nanobody – Cy5 (~10 μ g/mL), FITC-conjugated anti-myc antibody (1:10,000 dilution), or both BC2 nanobody-GFP alone in 500 μ L PBS-BSA and rotated at RT for 1 hour. After incubation, two final washes with cold PBS-BSA were made to remove any unbound material and samples were taken to flow cytometry for analysis.

Yeast: 50 mL culture of yeast displaying mSA2 with BC2T were grown in a 250 mL baffled flask containing SD-CAA for 2-3 days at 30 °C with shaking. After 2-3 days of growth in SD-CAA, the samples were subcultured in SD-CAA at an initial density of 1×10^7 cells/mL and grown to a density of $2-5 \times 10^7$ cells/mL. Yeast were subcultured again to a concentration of 1.0×10^7 cells/mL in SG-CAA (Galactose containing induction media) and grown for 2 days shaking at 250 rpm at a temperature of 20 °C. Approximately 10^8 cells were pelleted and washed with 500 μ L of 4 °C PBS-BSA. Yeast were subsequently incubated with either BC2 nanobody – Cy5 (~10 μ g/mL), FITC-conjugated anti-*myc* antibody / FITC conjugated anti-HA antibody (1:10,000 dilution), both the nanobody and one antibody, or BC2 nanobody – GFP (50 nM) alone in 500 μ L PBS-BSA and rotated at RT for 1 hour. After incubation, two final washes with cold PBS-BSA were made to remove any unbound material and samples were taken to flow cytometry analysis.

Western Blot Analysis

Using commercially available antibodies: Purified proteins were separated by SDS-PAGE and transferred to PVDF membrane via an iBlot Western blot apparatus. The membrane was blocked with 1X PBS, 5% milk, and 0.1% Tween-20 for 1 hour at RT. Primary antibodies for *myc* and HA tag were incubated separately with the appropriate membranes overnight at a 1:10,000 dilution in 10 mL of 1X PBS, 5% BSA, and 0.1% Tween-20 at 4 °C. Membranes were washed 3X with 1X PBS containing 0.1% Tween-20 and then incubated with Anti-Rabbit (Alexa Fluor 790) at a 1:10,000 dilution in 10 mL PBS, 5% milk, and 0.1% Tween-20 for 1 hour at RT. The membranes were then washed 3X with 1X PBS containing 0.1% Tween-20 and imaged in 1X PBS using the Odyssey Classic Infrared Imager.

Using “in-house” prepared nanobody-IR800 dye: Purified proteins were separated by SDS-PAGE and transferred to a PVDF membrane via an iBlot Western blot apparatus. The membrane

was incubated with 1X PBS, 5% milk, and 0.1% Tween-20 for 1 hour at RT. The BC2 nanobody-IR800 dye conjugate was then incubated overnight at ~0.10 μ M concentration in 10 mL of 1X PBS, 5% BSA, and 0.1% Tween-20 at 4 °C. The membrane was then washed 3X with 1X PBS containing 0.1% Tween-20 and imaged in 1X PBS using the Odyssey Classic Infrared Imager.

Appendix C.2: Chapter Five – Proteins Used in this Work

BC2 nb-Cys

MGMQVQLVESGGGLVQPGGSLTLSTASGFTLDHYDIGWFRQAPGKEREGVSCINN
SDDDTYYADSVKGRFTIFMNNAKDTVYLMNSLKPEDTAIYYCAEARGCKRGRYEYDF
WGQGTQVTVSSKKKHHHHHHC

BC2nb-nLuc

MGMQVQLVESGGGLVQPGGSLTLSTASGFTLDHYDIGWFRQAPGKEREGVSCINN
SDDDTYYADSVKGRFTIFMNNAKDTVYLMNSLKPEDTAIYYCAEARGCKRGRYEYDF
WGQGTQVTVSSKKKGGSMVFTLEDFVGDWRQTAGYNLDQVLEQGGVSSLFQNLGVSV
TPIQRIVLSGENGLKIDIHVIIPYEGLSGDQMGQIEKIFKVVYPVDDHHFKVILHYGTLVID
GVTPNMIDYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLINPDGSLLFRVTINGVTG
WRLCERILAHHHHHH

BC2nb-GFP

MQVQLVESGGGLVQPGGSLTLSTASGFTLDHYDIGWFRQAPGKEREGVSCINNSD
DDTTYADSVKGRFTIFMNNAKDTVYLMNSLKPEDTAIYYCAEARGCKRGRYEYDFW

GQGTQVTVSSKKKGGSGGSGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGE
GDATNGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGY
VQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNHNVYI
TADKQKNGIKANFKIRHNVEDGSGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDP
NEKRDHMLLEFVTAAGITHGMDELYKHHHHHH

GFP-HA

MAMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNHNHNVYIMADKQKNGIKVNFKIRHNI
EDGSGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGIT
LGMDELYKGGSGGSHHHHHHYPYDVPDYA

GFP-*myc*

MAMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNHNHNVYIMADKQKNGIKVNFKIRHNI
EDGSGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGIT
LGMDELYKGGSGGSHHHHHHEQKLISEEDL

GFP-BC2T

MEFMGVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTR

AEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNShNVYIMADKQKNGIKVNFKIR
HNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAA
GITLGMDELYKGGSGGSHHHHHHPDRKAAVSHWQQ

BC2T-mSA2-*myc* (*E. coli*)

MKKIACLSALAAVLAFTAGTSVAGQSGQAAAPGEQKLISEEDLGAPTYGASAEAGIT
GTWYNQHGSTFTVTAGADGNLTGQYENRAQGTGCQNSPYTLTGRYNGTKLEWRVIEW
NNSTENCHSRTEWRGQYQGGAEARINTQWNLTYEGGSGPATEQGQDTFTKVKPSAASG
SGGGSPDRKAAVSHWQQPRVGGGSGGGSGGGSGGGSGGGSGGGSGGQSGQSGDYNKNQYY
GITAGPAYRINDWASIYGVVGVGYGKFQTTEYPTYKHDTSDYGFSYGAGLQFNPMENV
ALDFSIEQSRIRSVDVGTWILSVGYRFGSKSRRATSTVTGGYAQSDAQGMNKMGGFN
LKRYEEDNSPLGVIGSFTYTEKSRTAS

HA-mSA2-BC2T-*myc* (yeast)

MQGVFEYYKSVTFVSNCGSHPSTTSKGSPINTQYVFKDNSSTIEGRYPYDVPDYALQ
ASGGGGSGGGGSGGGGSASEFASAEAGITGTWYNQHGSTFTVTAGADGNLTGQYENR
AQGTGCQNSPYTLTGRYNGTKLEWRVIEWNNSTENCHSRTEWRGQYQGGAEARINTQ
WNLTYEGGSGPATEQGQDTFTKVKPSAASGSGGSPDRKAAVSHWQQGSEQKLISEEDL

Appendix C.3: Chapter Five – Supplemental Information

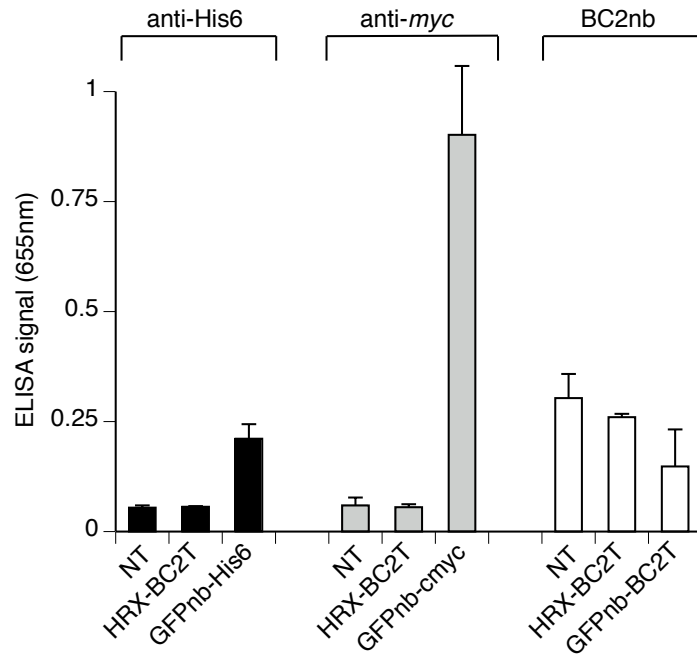


Figure S5. 1 ELISA-HRP. ELISA data using HRP conjugated antibodies or nanobodies for analysis. All wells were free (no immobilization on the plate's surface). To test off-target binding of anti-His6 antibody-HRP, anti-myc antibody-HRP, and BC2nb-HRP to non-immobilized well, wells were incubated with just buffer (NT), HRX-BC2T, the GFP nanobody with different tags depending on which antibody was used (anti-His6, colored black, anti-myc, colored grey, and BC2 nanobody-HRP, colored white) and washed to remove unbound material. After a 30-minute incubation with TMB-one substrate, plate was read at 655 nm. Signal is the observed absorbance at 655 nm. All experiments were performed in triplicate. Error bars represent the standard deviation of three experiments. NT = no treatment.

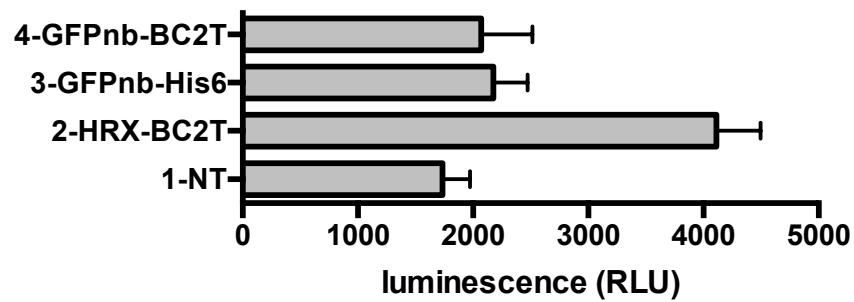


Figure S5. 2 ELISA – NanoLuciferase. ELISA data using fusion protein BC2nb-nLuc for analysis. All wells were without anything immobilized on the plate's surface. To test off-target binding of BC2nb-nLuc to non-immobilized wells, wells were incubated with just buffer (NT), HRX-BC2T, the GFP nanobody with either a His6x tag or BC2T and washed to remove all unbound material. Luminescence was read after a 10-minute incubation with NanoGlo substrate. All experiments were performed in triplicate. Error bars represent standard deviation of three experiments. NT = no treatment. RLU = relative luminescence units.

Table S5. 1 Flow Cytometry Data – Bacteria. Cy5, FITC, and GFP detected by flow cytometry to indicate display. All experiments were completed in triplicate. Values represent the mean of those experiments.

	construct	induced	incubated with	Cy5 (+)	FITC or GFP (+)
1	Bacteria - mSA2	-	-	0.87	1.73
2		Yes	-	1.72	1.23
3		-	Myc-ab-FITC	1.17	0.82
4		Yes	Myc-ab-FITC	3.11	94
5		-	BC2-nb Cy5	9.2	0.86
16		Yes	BC2-nb Cy5	98.6	1.9
7		-	BC2-nb-GFP	1.09	11.5
8		Yes	BC2-nb-GFP	3.68	98.1
9		Yes	BC2-nb Cy5 + Myc-ab- FITC	98.7	93.1

Table S5. 2 Flow Cytometry Data- Yeast. Cy5, FITC, and GFP detected by flow cytometry to indicate display. All experiments were performed in triplicate. Values represent the mean of those experiments.

	construct	induced	incubated with	Cy5 (+)	FITC or GFP (+)
1		-	-	0.63	6.11
2		Yes	-	0.17	1.1
3		-	Myc-ab-FITC	0.27	0.67
4		Yes	Myc-ab-FITC	0.53	71.7
5		-	HA-ab-FITC	0.28	0.71
6		Yes	HA-ab-FITC	0.58	69.3
7	Yeast - mSA2	-	BC2-nb Cy5	1.01	0.81
8		Yes	BC2-nb Cy5	71.1	1.61
9		-	BC2-nb-GFP	0.28	0.82
10		Yes	BC2-nb-GFP	0.57	59.6
11		Yes	BC2-nb Cy5 + Myc-ab-FITC	70.9	71
12		Yes	BC2-nb Cy5 + HA-ab-FITC	71.2	69.8

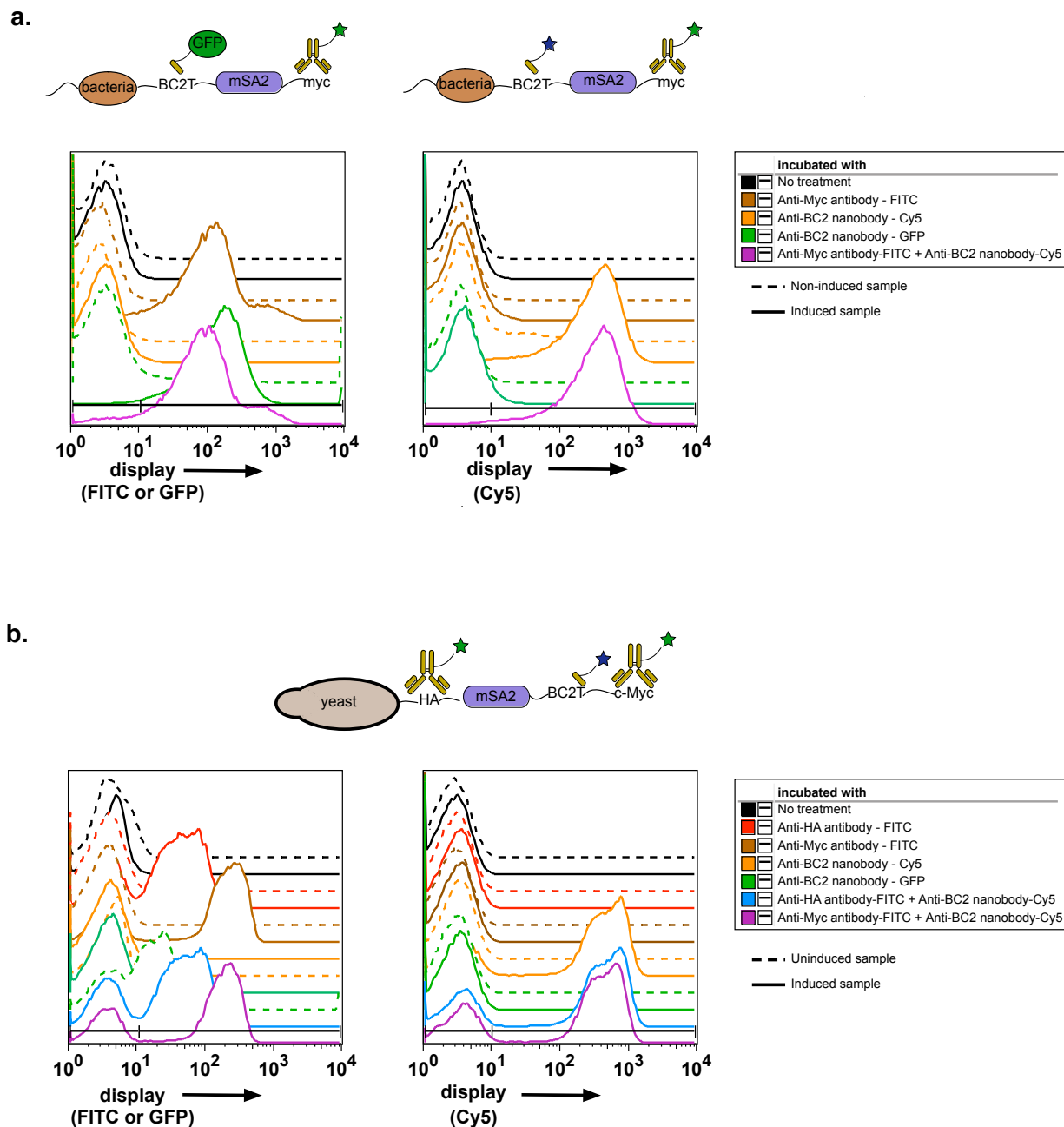


Figure S5. 3 Representative Flow Cytometry Histograms. Representative histogram of flow cytometry data for display of mSA2 on (a) bacteria and (b) yeast. In all cases, bacteria or yeast displaying mSA2 were detected after incubation with anti-HA antibody-FITC, anti-*myc* antibody-FITC, BC2 nanobody-Cy5, or BC2 nanobody-GFP. Non-induced samples are shown as dashed lines and induced samples as solid lines.

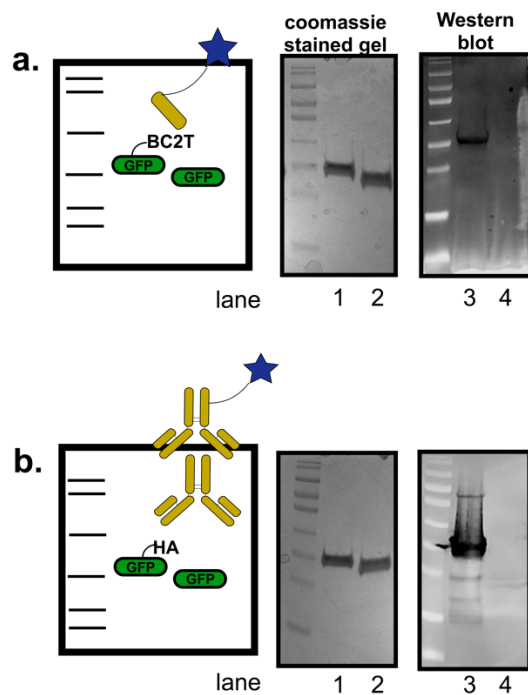


Figure S5. 4 Selectivity for Epitope Validated *via* Western blot. (a) 5 μ M coomassie stained gel and Western blot analysis of GFP-BC2T and GFP. Western blot analysis used BC2 nanobody-IRdye800. (b) 5 μ M coomassie stained gel and Western blot analysis of GFP-HA and GFP. Western blot analysis used anti-HA antibody and was visualized with Donkey Anti-Rabbit IgG Alexa Fluor 790.

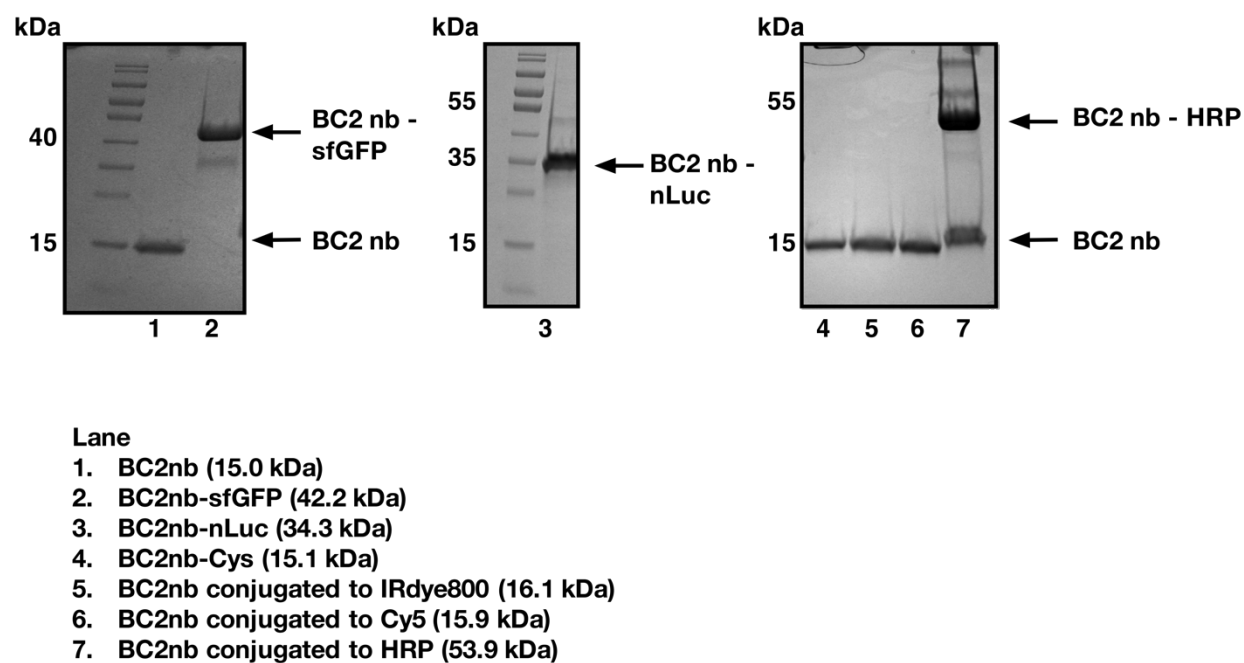


Figure S5. 5 SDS-PAGE of BC2 nanobodies utilized in this work. SDS-PAGE of purified BC2 nanobody and its fusion proteins and conjugates used in this work. BC2nb-HRP is the only construct to not receive complete conjugation, however because the assay is done in excess of reagent a significantly larger signal when compared to off-target controls was still achieved.

Appendix D: Supplemental Information of Chapter Six

Appendix D.1: Chapter Six – Part I - Methods

Materials

All chemicals obtained from Sigma-Aldrich unless specified

LB Miller Broth – Fisher

Phosphate buffered saline (PBS) – Corning Cell Grow

Lipofectamine 2000 – Life Technologies

0.25% Trypsin – Hyclone/Thermo Scientific

Fetal Bovine Serum (FBS) – Atlanta Biologicals

Dulbecco's Modified Eagle Medium (DMEM) – Hyclone/Thermo Scientific

Mammalian cell culture dishes – Fisher Scientific

Black Visiplate TC – PerkinElmer

RIPA Buffer – Boston BioProducts

5-alpha chemically competent *E. coli* – NEB

BL21 (DE3) chemically competent *E. coli* – NEB

Agar – GoldBio Technology

Carbenicillin – GoldBio Technology

Restriction Enzymes – NEB

Isopropyl- β -D-1-thiogalactopyranoside (IPTG) – GoldBio Technology

cComplete ULTRA Tablets, Mini, EDTA-free – Roche

Nano-Glo Live Cell Assay System – Promega

Quick Ligation Kit – NEB

Vent Polymerase – NEB

Oligonucleotides – IDT

Miniprep Kits – OMEGA

PageRuler Prestained Protein Ladder – Thermo Scientific

Ready Gel Precast Gels – BioRad

Hek293 Cells –ATCC

All water was obtained from a Milli-Q water purification system.

Instrumentation

Synergy Mx Microplate Reader – BioTek

EVOS FL Microscope – Advanced Microscopy Group

Sonifer W – 350 cell disrupter – Branson

MJ Mini Gradient Thermal Cycler – BioRad

Molecular Imager Gel Doc XR+ System – BioRad

Cloning

Purified proteins: All plasmids were constructed on a pETDuet-1 backbone. All proteins were assembled from a set of overlapping oligonucleotides or purchased g-block. Constructs were amplified using vent and then ligated into *NcoI* and *KpnI* restriction enzyme cleavage sites in the pETDuet-1 plasmid.

Mammalian expression: A mammalian GFP expression vector was generously provided by the Stasevich Lab (CSU). The gene coding for FP-LgBit was assembled from a set of overlapping

oligonucleotides and purchased g-block. FP-LgBit was amplified using vent and then ligated into *NheI* and *BglII* restriction enzyme cleavage sites in the mammalian GFP expression vector.

Protein Purification

Plasmids encoding GFP variants were transformed into BL21s (DE3) and plasmids encoding nanobody variants were transformed into Shuffle T7 Express cells. Cells were grown in either 1L or 500 mL LB cultures containing carbenicillin at 37 °C to OD₆₀₀ ≈ 0.5 and induced with 1 mM IPTG at 20 °C overnight. Cells were then collected by centrifugation and resuspended in phosphate buffer with 2 M NaCl (20 mM Sodium Phosphate, pH 7.4) and stored at -20 °C. Frozen pellets were thawed and incubated with cOmplete ULTRA protease inhibitor tablets then sonicated for 2 minutes. The lysate was cleared by centrifugation (8000 rpm; 20 minutes) and the supernatant was mixed with 1 mL Ni-NTA resin for 30 minutes. The resin was collected by centrifugation (4750 rpm, 10 minutes). The resin was washed with 50 mL buffer and 20 mM imidazole then 10 mL buffer with 50 mM imidazole. The protein was then eluted with 7 mL buffer containing 400 mM imidazole. The proteins were dialyzed against buffer with 150 mM NaCl and analyzed for purity by SDS-PAGE. Purified proteins were quantified using absorbance at 280 nm.

Mammalian Cell Culture

Hek293 cells were cultured in high glucose Dulbecco's modified Eagle medium (DMEM) with 10% FBS. Cells were incubated at 37 °C with 5% CO₂ environment. Cells were obtained from ATCC.

In Vitro Analysis

Individual aliquots of purified protein containing the SmBit tag were mixed at varying concentrations with excess purified FP-LgBit (1 μ M) in a black walled 96-well plate with shaking at RT for ~10 minutes. Next, a 1:50 dilution of NanoGlo reagent was added to wells and let continue shaking at RT for ~10 minutes before reading luminescence on a plate reader.

In Vivo Analysis

Hek293 cells were grown to ~85% confluency in a 24-well plate (clear bottom, black well). Prior to transfection, mammalian expression plasmid encoding FP-LgBit was ethanol precipitated. Cells were transfected with Lipofectamine 2000 following manufacturers protocol with a mammalian expression vector engineered to produce FP-LgBit. After ~10 hours of transfection, cells were washed 1X with OptiMem and then incubated with OptiMem solution containing purified proteins for a specified time. Cells were then washed 1X with OptiMem and 2X with PBS containing heparin sulfate 20 U/mL. Each round of incubation with heparin sulfate was allowed to incubate at 37 °C, 5% CO₂ environment for 10 minutes. Cells were then incubated with PBS and imaged on the EVOS FL fluorescence microscope to confirm transfection efficiency. After, cells were mixed with a 1:50 dilution of NanoGlo reagent and allowed to incubate with shaking at RT for ~10 min. Luminescence was measured on a Synergy Mx microplate reader.

Appendix D.2: Chapter Six – Part I: Proteins Used in this Work

sfGFP-LgBit

MVSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICTTGKLPVP
WPTLVTTLTLYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNFNHSHNVYITADKQKNGIKANFKIRHNVEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMLLEFVTAAGITHG
MDELYKSGLRSGGSGGSGGSGGVFTLEDVFGDWEQTAAYNLDQVLEQGGVSSLLQNL
AVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYPVDDHHFKVILPYG
TLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTLWNGNKIIDERLITPDGSMLFRVTIN
SHHHHHH

β -lac nanobody – 86

MAMAVQQLVESGGGSGVQAGGSLRLSCTASGGSEYSYSTFSLGWFRQAPGQEREAV
AAIASMGGLTYYADSVKGRFTISRDNANKNTVTLMNNLKPEDTAIYYCAAVRGYFMRL
PSSHNFRYWGGGTQVTVSSHHHHHHHGGGGSGGGGSGVSGWRLFKKIS

β -lac pcNanobody – 86

MAQVQLVEKGGGKVRAGGKLRLRCTASGGSEYSYSTFSLGWFRQAPGREREAVAA
IASMGGLTYYADSVKGRFKIKRDNANKNTVTLMNNLKPEDTAIYYCAAVRGYFMRLPS
SHNFRYWGGGTQVTVSRHHHHHHHGGGGSGGGGSGVSGWRLFKKIS

sfGFP-86

MASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICTTGKLPVP
WPTLVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNFNShNVYITADKQKNGIKANFKIRHNVEDG
SVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHG
MDELYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

Arginine Grafted GFP -86

MAMVSKGEELFTGVVPILVRLRGRVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKL
PVPWPTLVTTLTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAR
VKFEGDTLVNRIRLKGIDFKEDGNILGHKLEYNFNShNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGI
TLGMDELYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

+15GFP – 86

MAMGGASKGERLFTGVVPILVELDGDVNGHKFSVRGEGEGDATRGKLTCLKFICTTG
KLPVPWPTLVTTLTLYGVQCFSRYPKHKRHDFFKSAMPEGYVQERTISFKKDGTYKTR
AEVKFEGRTLNVNRIELKGRDFKEKGNILGHKLEYNFNShNVYITADKRKNGIKANFKIR
HNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSALSKDPKEKRDHMLLEFVTA
AGITHGMDELYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

+34GFP – 86

MGGASKGERLFRGKVPILVELKGDVNGHKFSVRGEGKGDATRGKLTCLKFICTTGKL
PVPWPVLVTTLTYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAE
VKFEGRTL VNRIKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHN
VKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMLLEFVTAAGI
KHGRDERYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

Aurein 1.2 – sfGFP – 86

MGGLFDIIKRIAESFGGSGGSGGSASKGEELFTGVVPILVELDGDVNGHKFSVRGEGE
GDATNGKLTCLKFICTTGKLPVPWPVLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGY
VQERTISFKDDGTYSKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYI
TADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDP
NEKRDHMLLEFVTAAGITHGMDELYKHHHHHHGGGGSVSGWRLFKKIS

Aurein 1.2 – β -lac nanobody – 86

MGGLFDIIKRIAESFGGSGGSGGSAQVQLVESGGGSVQAGGSLRLSCTASGGSEYSY
STFSLGWFRQAPGQEREAVAAIASMGGLTYADSVKGRFTISRDNKNTVTLMNNLK
PEDTAIYYCAAVRGYFMRLPSSHNFRYWGGGTQVTVSSHHHHHHGGGGSGGGGSVSG
WRLFKKIS

Aurein 1.2 - β -lac pcNanobody – 86

MGGLFDIIKRIAESFGGSGGSGGSAQVQLVEKGGGKVRAGGKLRLRCTASGGSEYS
YSTFSLGWFRQAPGREREAVAAIASMGGLTYADSVKGRFKIKRDNAKNTVTLRMNNL
KPEDTAIYYCAAVRGYFMRLPSSHNFRYWGQGTRVTVSRHHHHHHGGGGSGGGGSVS
GWRLFKKIS

Aurein 1.2 Arginine Grafted GFP-86

MGGLFDIIKRIAESFGGSGGSGGSASKGEELFTGVVPILVRLRGRVNGHKFSVSGEGE
GDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGY
VQERTIFFKDDGNYKTRARVKFEGDTLVNRIRLKGIDFKEDGNILGHKLEYNYNSHNVY
IMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKD
PNEKRDHMLLEFVTAAGITLGMDELYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

Aurein 1.2 +15GFP-86

MGGLFDIIKRIAESFGGSGGSGGSASKGERLFTGVVPILVELDGDVNGHKFSVRGEGE
GDATRGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPKHKMRHDFFKSAMPEGY
VQERTISFKKDGTYKTRAEVKFEGRTLVRNRIELKGRDFKEKGNILGHKLEYNFNHSHNVYI
TADKRKNGIKANFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSALS KDP
KEKRDHMLLEFVTAAGITHGMDELYKHHHHHHGGGGSGGGGSVSGWRLFKKIS

Aurein 1.2 +34GFP-86

MGGLFDIIKRIAESFGGSGGSGGSASKGERLFRGKVPILVELKGDVNGHKFSVRGEG
KGDATRGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPKHKMRHDFFKSAMPKG

YVQERTISFKKDGKYKTRAEVKFEGRTL VNRIKLKGRDFKEKGNILGHKLRYNFNSHKV
 YITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSKLSK
 DPKEKRDHMLLEFVTAAGIKHGRDERYKHHHHHHGGGGSVSGWRLFKKIS

Appendix D.3: Chapter Six – Part I: Supplemental Information

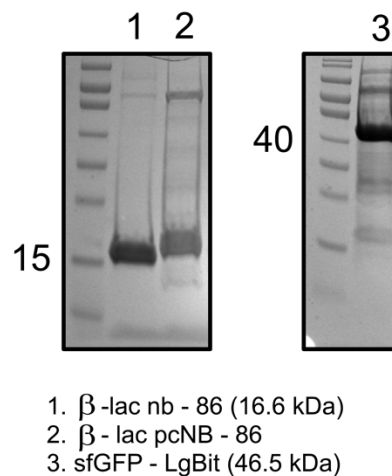
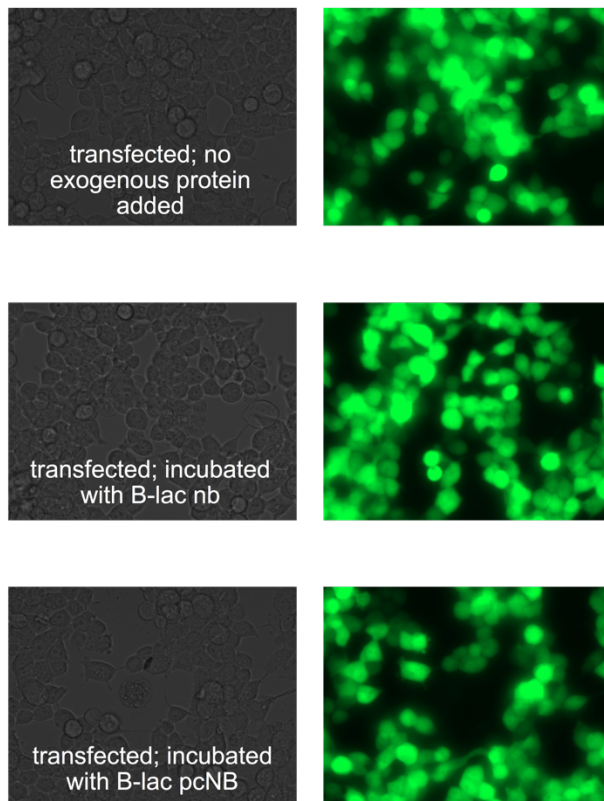


Figure S6. 1 SDS-PAGE of Purified Proteins – NanoBit. PAGE analysis of purified wildtype and polycationic resurfaced nanobodies with SmBit tag and purification of sfGFP-LgBit.

24- well plate - 10-12 hours after transfection
Hek293 Cells



All transfected with 1 μ L Lipofectamine 2000 and 150
ng DNA

Figure S6. 2 Live-Cell Fluorescent Microscopy Images – NanoBit. Cells were transfected with plasmid encoding sfGFP-LgBit. Prior to luminescence assay, microscopy images were used to comment on transfection efficiency.

Appendix D.4: Chapter Six – Part II - Methods

Materials

All chemicals were obtained from Sigma-Aldrich unless specified

Phosphate buffered saline (PBS) – Hyclone/Thermo Scientific

Brilliant Blue R-250 – J.T. Baker

Imidazole – Sigma Aldrich

PageRuler Prestained Protein Ladder – Thermo/Scientific

LB Miller Broth – Fisher

5 – alpha chemically competent *E. coli* – NEB

BL21 (DE3) chemically competent *E. coli* – NEB

Agarose A – Bio Basic Inc.

Carbenicillin – GoldBio Technology

Kanamycin – GoldBio Technology

Restriction Enzymes – NEB

Isopropyl - β -D-1 thioglactopyranoside (IPTG) – GoldBio Technology

cComplete ULTRA Tablets, Mini, EDTA-free – Roche

Quick Ligation Kit – NEB

Vent Polymerase – NEB

Oligonucleotides/gblocks – IDT

Miniprep Kits – Omega

12% Ready Gel Precast Gels – BioRad

Snakeskin Dialysis Tubing 10 K MWCO – Thermo Scientific

All water obtained from Milli-Q water purification system

Instrumentation

Sonifer W – 350 cell disrupter – Branson

MJ mini gradient thermal cycler – BioRad

Molecular Imager Gel Doc XR+ System – BioRad

NanoDrop 200 UV-Vis Spectrophotometer – Thermo Scientific

Stasevich Laboratory Custom-Built Microscope

Microscopy

All cell images and videos were taken on the Stasevich Laboratory Microscopy, a custom-built widefield fluorescence microscope with a highly inclined illumination scheme.¹

Mammalian Cell Culture

U2OS cells were prepared in the Stasevich Laboratory and cultured in DMEM media. Cells were incubated at 37 °C with 5% CO₂ environment.

Cloning

Purified proteins: All plasmids were constructed on a pETDuet-1 backbone. All proteins were assembled from a set of overlapping oligonucleotides or purchased g-block. Constructs were amplified using vent and then ligated into *NcoI* and *KpnI* or *BamHI* and *PacI* restriction enzyme cleavage sites in the pETDuet-1 plasmid.

Mammalian expression vector: A vector encoding 10XFLAG-tag (SM) beta-actin cut with restriction enzymes *AgeI* and *ClaI* was generously given by the Stasevich Laboratory. A 7X TAR DNA fragment was obtained as a g-block from IDT with flanking *AgeI* and *ClaI* cut sites.

Construct was amplified using vent and then ligated into *AgeI* and *ClaI* restriction enzyme cleavage sites in the mammalian expression vector.

A vector encoding sfGFP-H2B was modified to make Cpep-H2B.

Protein Purification

Plasmids were transformed into BL21s (DE3). Cells were grown in either 2500 mL or 500 mL LB cultures containing carbenicillin at 37 °C to an $OD_{600} \approx 0.5$ and induced with 1 mM IPTG at 20 °C overnight. Cells were then collected by centrifugation and resuspended in phosphate buffer with 2 M NaCl (20 mM Sodium Phosphate, pH 7.4) and stored at -20 °C. Frozen pellets were thawed and incubated with cOmplete ULTRA protease inhibitor tablets and then sonicated for 2 minutes. The lysate was cleared by centrifugation (8000 rpm, 20 minutes) and the supernatant was mixed with 1 mL Ni-NTA resin for 30 minutes. The resin was collected by centrifugation (4750 rpm, 10 minutes). The resin was washed with 50 mL buffer and 20 mM imidazole then 10 mL buffer and 50 mM imidazole. The protein was then eluted with 7 mL buffer containing 400 mM imidazole. The proteins were dialyzed against buffer with 150 mM NaCl and analyzed for purity by SDS-PAGE. Purified proteins were quantified using absorbance at 280 nm.

Appendix D.5: Chapter Six – Part II: Proteins Used in this Work

sfGFP-TBP

MGSSHHHHHSQDPMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGK
LTLKFICTTGKLPVPWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTISF
KDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNShNVYITADKQKN

GIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH
MVLLEFVTAAGITHGMDELYKGGGGSGGGGSMAQVQLQVDMAPETRPNHTIYINNL
NSKIKKDELKKSLEYAIFSQFGQILDILVPRQRTPRGQAFVIFKEVSSATNALRSMQGYPFY
DKPMRIQYARTDKRIPAKMKGT

sfGFP-5helix

MGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTCLKFICTTGKLPV
PWPTLVTTLTYGVCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVK
FEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVED
GSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH MVLLEFVTAAGITHG
MDELYKSGLRSGGSGGGSGGGSTQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQ
ARILAGGSGGHTTWMEWDREINNYTSLIHS LIEESQNQQEKNEQELLEGGSSGGQLLSGIV
QQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAGGSGGHTTWMEWDREINNYTSLIHS
LIEESQNQQEKNEQELLEGGSSGGQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQA
RILAGGHHHHHH

C-peptide

WMEWDREINNYTSLIHS LIEESQNQQEKNEQELL

7X TAR DNA

ACAACCGGTAACCTACAAAAGGCCAGATCTGAGCCTGGGAGCTCTCTGGCCACG
TTATGTACGTTGCTCCGTCAAAGGCGCGCGCAGATCTGAGCCTCCGAGCTCTCTGCG
CACCTTTAACCCTCCTGAGAACCGGGAGGCGGGAATCCGTCACGCAGATCGAGCCT

GGGAGCTCTCTGCGGGGGCACTGGCAACCAACTTCTCGGGTCCTGCCCCGCCGGAGA
TCTGAGCCTGGGAGCTCTCTCCGGTAAGTTTTTCCACTCGCTTGAGCCGGCTAGGCA
GATCGAGCCTGGGAGCTCTCTGCCATTCAGGATCACGTTACCGCCAAAAAATGGGA
CCGGAGATCTGAGCCTCCGAGCTCTCTCCGGAGTTAGGCATAAGGCTGCATGCTACC
TTGTCGCGAGATCTGAGCCTGGGAGCTCTCTATGCACATCAAACACCTCAGATCTCA
TATCGATAG (*AgeI* to *ClaI* restriction enzyme sites)

Appendix D.6: Chapter Six – Part II: Supplemental Information

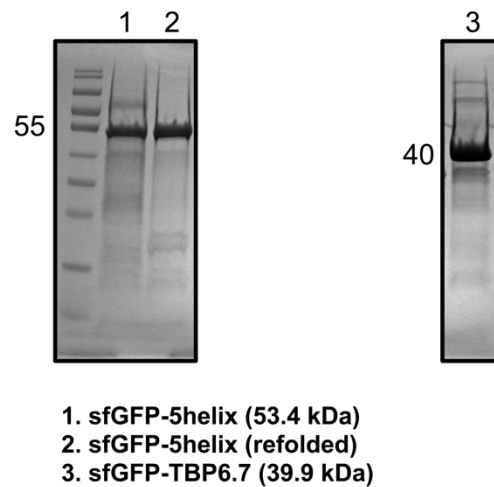


Figure S6. 3 SDS-PAGE of Proteins Used in this Work – Imaging Constructs.

ABBREVIATIONS

ADCC	antibody-dependent cellular cytotoxicity
Ala/A	alanine
Antp	antennapedia-homeodomain-derived peptide
Arg/R	arginine
Asn/N	aspartic acid
CD	circular dichroism
CDC	complement-dependent cytotoxicity
CDR	complementary determining region
CPP	cell-penetrating peptide
Cy5	cyanine dye number 5
Cys/C	cysteine
Da	Dalton
DAPI	4'6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
ELISA	enzyme-linked immunosorbent assay
FACS	fluorescence-activated cell sorting
FITC	fluorescein isothiocyanate
GFP	green fluorescent protein
Gln/Q	glutamine
Glu/E	glutamic acid
Gly/G	glycine

HER2	human epidermal growth factor receptor 2
His/H	histidine
HIV	human immunodeficiency virus
HRP	horseradish peroxidase
IgG	immunoglobulin G
Ile/I	isoleucine
IPTG	β -D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
K _D	dissociation constant
Leu/L	leucine
Lys/K	lysine
MCS	multiple cloning site
Met/M	methionine
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
mRNA	messenger RNA
MW	molecular weight
NHR	N-heptad repeat
nLuc	nano-luciferase
nt	nucleotides
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein data bank

Pen	penetratin
Phe/F	phenylalanine
PSMA	prostate specific membrane antigen
PPI	protein-protein interaction
Pro/P	proline
PTD	protein transduction domain
RNA	ribonucleic acid
RT	room temperature
RPM	revolutions per minute
SAV	streptavidin
SDS	sodium dodecyl sulfate
Ser/S	serine
siRNA	small interfering ribonucleic acid
TAT	trans-activating transcriptional activator
TAR	trans-activating response element
Thr/T	threonine
TMB	3,3',5,5'-tetramethylbenzidine
Trp/W	tryptophan
Tyr/Y	tyrosine
U1A	human U1 snRNP protein
U1hpII	U1 hairpin II of the human snRNP complex
Val/V	valine
WB	Western blot